

THE TWO-ENVELOPE PROBLEM: AN INFORMED CHOICE

By  
JEFFREY BRIAN TYLER

at the  
UNIVERSITY OF SOUTH AFRICA

submitted in accordance with the requirements for the degree of

MASTER OF SCIENCE

in the subject

OPERATIONS RESEARCH

at the  
UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF P H POTGIETER

CO-SUPERVISOR: PROF G DAVIE

MARCH 2014

Mathematics is the DNA of all that matters.

## Acknowledgements

Professor Petrus Potgieter and Dr George Davie never gave up. In all my thinking and doing, Petrus and George believed in me and gave me every opportunity to realise my long held dream. Their support and understanding went beyond the academic realm, and for that, I am deeply grateful. Their unwavering attention, guidance, and enthusiasm sustained me through an arduous journey and I could not have asked for better companions. Petrus and George, I thank you.

# Abstract

The host of a game presents two indistinguishable envelopes to an agent. One of the envelopes is randomly selected and allocated to the agent. The agent is informed that the monetary content of one of the envelopes is twice that of the other. The dilemma is under which conditions it would be beneficial to switch the allocated envelope for the complementary one. The objective of his or her envelope-switching strategy is to determine the benefit of switching the allocated envelope and its content for the expected content of the complementary envelope.

The agent, upon revealing the content of the allocated envelope, must consider the events that are likely to have taken place as a result of the host's activities. The preceding approach is in stark contrast to considering the agent's reasoning for a particular outcome that seeks to derive a strategy based on the relative contents of the presented envelopes. However, it is the former reasoning that seeks to identify what the initial amounts could have been, as a result of the observed amount, that facilitates the identification of an appropriate switching strategy.

Knowledge of the content and allocation process is essential for the agent to derive a successful switching strategy, as is the distribution function from which the host sampled the initial amount that is assigned to the first envelope.

For every play of the game, once the agent is afforded the opportunity of sighting the content of the randomly allocated envelope, he or she can determine the expected benefit of switching.

# Contents

<b>Contents</b>	<b>v</b>
<b>Nomenclature</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Information . . . . .	3
1.1.1 Content and allocation . . . . .	4
1.1.2 The rules of the game . . . . .	7
1.2 Expectation . . . . .	8
1.3 A strategy or a paradox . . . . .	9
1.4 Summary . . . . .	11
<b>2 Literature Review</b>	<b>14</b>
2.1 Posing the problem . . . . .	16
2.1.1 The formulation . . . . .	17
2.1.2 The content . . . . .	18
2.2 The method . . . . .	20
2.2.1 Reason and logic . . . . .	21
2.2.2 Probability . . . . .	21
2.3 Switching strategies . . . . .	23
2.3.1 The paradox . . . . .	24
2.4 Summary . . . . .	25
<b>3 Posing the two-envelope problem</b>	<b>26</b>
3.1 The host's activities . . . . .	27
3.2 The agent . . . . .	29
3.3 Envelope contents . . . . .	30
3.3.1 Doubling-only . . . . .	31

3.3.2	Halving-only . . . . .	31
3.4	The envelope allocation . . . . .	32
3.5	The benefit of switching envelopes . . . . .	33
3.6	Summary . . . . .	34
<b>4</b>	<b>The expected benefit of information</b>	<b>36</b>
4.1	Essential information for the agent . . . . .	37
4.2	The events associated with the sighted amount . . . . .	38
4.2.1	A “doubling-only” strategy . . . . .	39
4.2.2	A “halving-only” strategy . . . . .	41
4.2.3	A “halving or doubling” strategy . . . . .	43
4.2.4	Expected benefit associated with discrete variables . . . . .	43
4.3	Continuous random variables . . . . .	45
4.3.1	Probability . . . . .	46
4.3.2	Expectation . . . . .	47
4.3.3	A “doubling-only” strategy . . . . .	47
4.3.4	A “halving-only” strategy . . . . .	48
4.3.5	A “halving or doubling” strategy . . . . .	49
4.4	A switching strategy . . . . .	50
4.5	Other strategies . . . . .	51
4.5.1	Allocate the first envelope . . . . .	52
4.5.2	Allocate the second envelope . . . . .	53
4.6	Summary . . . . .	55
<b>5</b>	<b>Envelope-switching strategies</b>	<b>57</b>
5.1	Uniform distribution . . . . .	57
5.2	Weibull distribution . . . . .	58
5.2.1	Strategy associated with the “doubling-only” process . . . . .	58
5.2.2	Strategy associated with the “halving and doubling” process . . . . .	59
5.3	Broome’s approach . . . . .	61
5.3.1	Discrete variable . . . . .	61
5.3.2	Continuous variable . . . . .	61
5.4	Extreme values . . . . .	62
5.4.1	Discrete variable . . . . .	62
5.4.2	Continuous variable . . . . .	63
5.5	Recurrence relation . . . . .	66



5.6	Improper distribution . . . . .	67
5.6.1	Exponential . . . . .	67
5.6.2	Jeffreys' prior . . . . .	68
5.7	Summary . . . . .	70
<b>6</b>	<b>The question of a “paradox”</b>	<b>71</b>
6.1	A likely outcome . . . . .	72
6.2	The Story . . . . .	74
6.3	Half the story . . . . .	75
6.4	The full story . . . . .	77
6.5	Summary . . . . .	78
<b>7</b>	<b>Conclusion</b>	<b>80</b>
7.1	In closing . . . . .	82
7.2	The paradox . . . . .	83
	<b>Reference</b>	<b>85</b>

# Nomenclature

$X_1$	The random variable that represents the initial amount of money assigned to the first envelope
$X'_1$	The random variable that represents the amount of money assigned to the second envelope
$X_2$	The random variable that represents the toss of an unbiased coin
$X_3$	The random variable that represents the toss of an unbiased coin
$Y$	The random variable that represents the amount of money $y$ in the randomly selected and then allocated to the agent (player of the game)
$Z$	The random variable that represents the amount of money in the complementary envelope
<b>A</b>	The allocation matrix
<b>C</b>	The matrix representing the contents of the allocated and complementary envelopes
$B$	The random variable that represents the benefit (difference) between the content $z$ of the complementary envelope $Z$ and the content $y$ of the allocated envelope $Y$
$E(B Y = y)$	The expected benefit when switching the amount $Y = y$ in the allocated envelope for the expected content of the complementary envelope
$(\Omega, \mathcal{F}, P)$	The probability space

$\Omega$	The sample space
$\mathcal{F}$	The event space
$\mathcal{C}_y$	The set of events for consideration, by the agent, when observing the amount $y$ in the allocated envelope $Y$
$C_y$	A subset of $\mathcal{C}_y$
$c$	The event $c \in C_y$
$b(c)$	The benefit of switching for the event $c$
$e(y)$	The exchange condition necessary for switching the allocated envelope for the complementary envelope
$P(x)$	Equivalently, $P(X = x)$
$F_X(x)$	The cumulative distribution function of the random variable $X$
$f_X(x)$	The probability density function associated with $F_X(x)$
$X \sim f_X(x)$	The random variable $X$ is distributed according to the probability density function $f_X(x)$
$s(y)$	The strategy associated with the observed amount $y$
$\mathcal{G}$	The sub- $\sigma$ -algebra induced by $Y$
$g(\omega)$	A function that defines the factor applied to the initial amount to determine the amount assigned to the second envelope

# Chapter 1

## Introduction

... statistical information on frequencies within a large, repetitive class of events is strictly irrelevant to a decision whose outcome depends on a single trial.

— Risk, Ambiguity, and the Savage Axioms (Ellsberg 1961:2)

The two-envelope problem associated with the game of switching envelopes, which is responsible for the so-called “two-envelope paradox”, relies on the choices that confront an agent who must derive a strategy for retaining the randomly allocated envelope or switching envelopes and accepting the complementary envelope.

The simplicity of the game belies the unprecedented amount of scholarly interest and attention it has garnered since first considered by Kraitichik (1942:p133). A number of papers cite both the origination of the problem and the related paradox in a variant of the “necktie” problem first posed by Kraitichik (1942:p133). The game has since mutated into the current two-envelope problem, which has enjoyed notoriety because of its paradoxical interpretation in that it is alleged that it is always beneficial to switch envelopes irrespective of the content.

The “two-envelope paradox”, which has been the cause of much debate and consternation within the academic fraternity, centres on the strategy of an agent (player of the game) and rarely the related activities of the host (of the game).

While the two-envelope problem that has given rise to the paradox is reasonably simple in its presentation, its resolution and explanation have been somewhat elusive and, in some cases, unconvincing because most authors have failed to formulate the problem correctly. What is important, however, is that the agent can use the information he or she possesses in order to adopt a game strategy that allows him or her to determine the expected content of the complementary envelope upon sighting the content of the allocated envelope.

The objective of any strategy is to identify the benefit of switching the randomly allocated envelope and its observed contents for the expected content of the complementary envelope. It is very tempting to claim that the relative and therefore average content of the complementary envelope must be greater than the observed amount in the randomly allocated envelope for switching to take place.

Of the two primed envelopes, one is randomly selected and allocated to the agent. The agent, who has been informed that the content of one of the envelopes is twice that of the other, is then given the option of retaining the allocated envelope or switching envelopes in order to secure the envelope that contains the larger amount.

In addition to being informed that the content of one of the envelopes is twice the content of the other, the agent must also consider whether having sight of the content of the allocated envelope is relevant. If sighting the content is deemed to be significant, then it becomes important to determine what other information would assist the agent in identifying the game strategy that can realise the greatest expected benefit for any observed amount in the allocated envelope. Although “twice the other” seems to be a simple statement, this result can derive from halving, doubling, or a combination of halving and doubling, depending on the content and allocation process.

In most previous articles, the authors assume that the agent will reason that if the randomly allocated envelope is observed to contain an amount of money, then the com-

plementary envelope must contain either half or double that amount, each with an equal likelihood. However, it is not the content of the complementary envelope that is relevant. It will be realised that it is the content of the first envelope and the consequent content of the second envelope that are directly related to the observed amount in the allocated envelope. And, by inference, indirectly the content of the complementary envelope is considered relevant, when in reality it is not.

## 1.1 Information

The host of the game is responsible for assigning amounts of money, according to a certain content and allocation process, to each of the two envelopes. How these amounts are identified and assigned to the envelopes is explained in the content and allocation process of the game. The money assigned to the first envelope is a randomly determined amount, whereas the second amount is dependent upon the amount assigned to the first envelope.

In the random allocation of these indistinguishable envelopes, the agent receives one of the two envelopes.

Newcomers to the two-envelope problem may be forgiven for an initial response that has been exhaustively researched and, depending on the objective of the study in question, justified. In most instances, such examinations have ignored the activities of the host — the identification of the initial amount that is to be assigned to the first envelope and the subsequent amount assigned to the second envelope — for a specific play of the game, with the presented envelopes and their contents treated as the only relevant data for the agent. Indeed, it has been assumed that the outcomes of the host's activities for each play of the game are the only factors deemed to be relevant for solving the problem.

If the agent considers the problem from a probabilistic perspective, it soon becomes

apparent that the activities of the host — for all related outcomes — directly influences the agent’s potential strategy. It also becomes clear that the observed amount may be related to the activities of the host on an expected basis and that there is an indirect relationship between the amounts outside of the immediately obvious linkage. Specifically, what the agent reveals in the allocated envelope does define the possible activities of the host (who was responsible for the initial amount assigned to the first envelope and consequent amount assigned to the second envelope).

This problem can be considered to be a two-stage experiment. The first stage is associated with the activities of the host, while the second stage is associated with the activities of the agent. Because these two stages are subtly independent, the agent must consider the following five content and allocation possibilities (one of which will have been adopted by the host) in order to derive his or her switching strategy.

### 1.1.1 Content and allocation

The initial amount is determined by the host who randomly samples from an appropriate cumulative distribution function and assigns the amount to the first envelope. Thereafter, the second amount is derived according to one of the following five possibilities:

**Possibility 1.** *The host may prime the second envelope, by halving or doubling the initial amount of money, and assigning the resultant amount to the second envelope. He or she then randomly selects one of the envelopes and allocates it to the agent for consideration;*

**Possibility 2.** *The host may prime the second envelope, by doubling the content of the first. He or she then randomly allocates one of the envelopes to the agent for consideration;*

**Possibility 3.** *The host may prime the second envelope, by halving the content of the first. He or she then randomly allocates one of the envelopes to the agent for consideration;*

While it is true that the content of the complementary envelope is either half or double the content of the randomly allocated envelope, it is true that the expected benefit of switching envelopes is the difference between the expectation of the content of the complementary envelope and the observed content of the allocated envelope.

In order to identify an appropriate switching strategy, it is thus relevant for the agent to consider what the host may have assigned initially to the first envelope and subsequently to the second envelope. The expected benefit of switching, which is a function of the amount observed by the agent, is associated with the host having assigned either half the amount, the equivalent amount, or twice the observed amount to the first envelope. And, the content of the second envelope is now considered as a function of the content of the first envelope. This, allocation of the content to the first and envelopes is not the end of the story. The envelopes are indistinguishable and subject to a random allocation that will see different combinations that ultimately generate the allocated and complementary envelopes. Since the host is randomly selecting an initial amount from an arbitrary distribution function it is distinctly possible that each of the amounts – half, equivalent, or double – will each occur with a different probability. Moreover, it is accepted that one of the indistinguishable envelopes is randomly allocated to the agent and that the complementary envelope will be considered to be preferred or not upon sighting the content of the allocated envelope, when permitted.

It is the random allocation of the envelopes that ensures that the allocated envelope contains a fixed amount and that the complementary envelope will contain either half or double the content of the allocated envelope. That the content of the complementary envelope contains half or double that of the allocated envelope agrees with the relative explanation that is so often quoted.

**Possibility 4.** *The host may allocate the first envelope to the agent and then prime the second envelope by either halving or doubling the content of the first; or*



**Possibility 5.** *The host may retain the first envelope. The second envelope will then be primed by either halving or doubling the content of the first envelope. This second envelope is then allocated to the agent.*

The traditional approach to solving the two-envelope problem has only considered the content  $Y$  of the allocated and the content  $Z$  of the complementary envelope. For any play of the game it will be taken that the contents of the envelopes remain unchanged. With this condition it must be accepted that, irrespective of the content and allocation process, that the relative content of the complementary envelope is either twice or half that of allocated envelope. For any specific play of the game (single trial) the allocated envelope may contain the larger amount  $Y = 2x$  and the complementary envelope  $Z = x$ , and as a result  $Z = Y/2$ . Alternatively, if the allocated envelope contains the smaller amount, then  $Y = x$  and  $Z = 2x$ , and as a result  $Z = 2Y$ . Which transformation should be applied is clearly an issue that cannot be ignored. In the absence of information the agent can conclude that the host primed the envelopes with the amounts  $X_1 = y$  and  $X'_1 = 2y$ , if the “doubling-only” process was adopted. Therefore, as a result of the random allocation, the agent may observe  $Y = y$  and consequently conclude that the complementary envelope contains  $Z = 2y$  or the agent may observe  $Y = 2y$  and consequently conclude that the complementary envelope contains  $Z = y$ . Since these outcomes are equally likely – due to symmetry – the agent should be indifferent to switching envelopes or not.

It would be more appropriate to take the content of the allocated envelope  $Y = y$  as fixed and to then consider the possible values that the complementary envelope  $Z$  may contain as a result of the content and allocation process. While it is acknowledged that the relative content of the complementary envelope  $Z$  can be either  $y/2$  or  $2y$ , these possible values should only be considered as a result of the content and allocation process.

The amount in the complementary envelope is assumed to be associated with the host generating an amount  $X_1 = y$  and does not recognise that the amount  $Y = y$  can only

be observed as a result of the host's activities associated with  $x_1 \in \{y/2; y; 2y\}$ . Further, the assertion that  $Y = y$  is only possible for  $X_1 = 2y$  or  $X_1 = y/2$  does not consider the relationship between the observed  $Y = y$  and  $X_1$  and states that this relationship is distribution free by having originated from the host generating  $X_1 = 2x_1$  or  $X_1 = x_1/2$ . The realisation that  $X_1 = 2y$  and  $X_1 = y/2$  are just as relevant as  $X_1 = y$  for the derivation of the expected benefit  $B$  is discussed in Chapter 4, where the alternative strategies associated with the content and allocation process are derived. The evolution of the two-envelope problem and the consequent paradox is then discussed in Chapter 6. The way in which the problem is posed lends itself to assumptions about the reasoning of the agent. That the agent is assumed to reason in a particular way is not necessarily correct. However, it does seem to be the approach followed by previous authors (Sobel 1994).

### 1.1.2 The rules of the game

While little or no attention has been paid to how the activities of the host influences the process of assigning the amounts to the envelopes and the subsequent allocation of the envelopes, it is important for the agent to know about these. This information must include the distribution function from which the host sampled the initial amount, while the agent must also have sight of the content of the allocated envelope.

The rules of the game for the informed agent are thus as follows:

**Rule 1.** *The agent will have sight of the contents in the allocated envelope;*

**Rule 2.** *The agent will be informed about the content and allocation strategy;*

**Rule 3.** *The agent will be informed about the process of amount identification;*

**Rule 4.** *The agent will be informed about the distribution function from which the host sampled the initial amount. If the agent is not informed about the distribution function*

*from which the initial amount is generated, the agent may assume a prior distribution. In addition, subsequent to each play of the game, the agent may refine the prior distribution to enhance the predictive ability in the successive play of the game. Irrespective of whether the agent has assumed or derived such a prior distribution or not, the density function under consideration for each play of the game will be taken as given;*

**Rule 5.** *The agent will be informed of the bounds (if they exist) for the amounts that can be assigned to the envelopes. While it is acknowledged that the paradox does not exist in the presence of a defined upper bound this still does not address a playing strategy.*

The relevance of each of these rules is fundamental for solving the two-envelope problem.

## 1.2 Expectation

The scholarly interest that the two-envelope paradox has generated is based on its simple presentation. But, it is the formulation of the problem (where the absence of information and as a result assumptions) that has perpetuated the assertion of the paradox.

The traditional explanation of the paradox is founded on the premise that the outcome of any game is the difference between what is expected and what is observed. In particular, what is expected is taken as the average of two amounts, namely (i) double the observed amount and (ii) half the observed amount.

As described earlier, the host of the two-envelope game prepares two indistinguishable envelopes by assigning amounts of money to each. The agent is randomly allocated one of the envelopes. Before any interaction with the envelopes, the agent is informed that the amount of money in one envelope is twice the amount in the other. The agent is then asked whether in the absence of sighting the content, the complementary envelope would be preferred.

However, nothing has been said about sighting the content. Herein, it is shown that in the absence of the distribution function from which the host identified the initial amount, there is no benefit for the agent to switching envelopes. It is only once the density function is considered that the agent can derive a switching strategy.

For many the approach is associated with the host will double the initial amount and assign this to the second envelope. After randomly allocating one of the indistinguishable envelopes to the agent, the agent can conclude that the envelope will contain an amount that could be the larger or the smaller. If the agent is informed that the host has adopted a doubling only process, then the agent can safely assume that if the envelope contains the larger amount then the initial amount assigned to the first envelope would have been the smaller amount — a relative amount that is also stated as half the amount in the allocated envelope. The probability of the host generating this smaller amount would have been identified from the distribution function chosen by the host, and this smaller amount would have been doubled, hence the larger amount. Alternatively, if the envelope contains the smaller amount then the agent can conclude that the initial amount was assigned to the first envelope and that the larger (double the smaller) was assigned to the second envelope — a relative amount that is also stated as twice the amount in the allocated envelope. The complementary envelope does, as noted by many, contain either half or double the amount in the allocated envelope. The content of the allocated envelope is fixed. Applying a probability to the respective initial amounts does yield the expected outcome and by subtracting the amount contained in the allocated envelope the expected benefit of switching can be derived.

### **1.3 A strategy or a paradox**

Although two envelopes are under consideration in every play of the game, only the content of the randomly allocated envelope is relevant once its content has been revealed for

any play. The content of the complementary envelope is irrelevant for the purposes of deriving a strategy unless the agent has not been given information of the prior distribution function from which the host is sampling (in which case the content of the second envelope can be utilised for updating the prior distribution). Hence, the agent can collate the outcomes for each play of the game in order to monitor the overall performance of the adopted strategy.

The traditional posing of the problem has led to it being improperly formulated and consequently incorrectly solve (Silver 1994). This ill-stated problem, in the absence of any assumption and information, together with the failure to re-formulate the problem correctly is the reason for the plethora of papers that continue to sustain the myth of the paradox (Loredo 2004).

The argument, for the envelopes containing  $X_1$  and  $X'_1$ , in the absence of sighting the content, is that if the agent had been randomly allocated the envelope that contained the amount  $Y = X_1$  then the benefit  $B$  of switching would be the difference between  $X'_1 - X_1$ . Alternatively, if the allocated envelope contained the amount  $Y = X'_1$ , then the benefit  $B$  of switching would be the difference between  $X_1 - X'_1$ . Each of the envelopes is equally likely to be allocated to the agent. The agent may assume that since the contents are fixed at the time of selection, they remain unchanged during the process; therefore, it is irrelevant whether he or she chooses to switch or not.

The alternative argument, in the absence of sighting the content, runs as follows. The complementary envelope contains either half or double the amount in the randomly allocated envelope for the following reason. If the content of the allocated envelope were observed to be  $Y$ , then the complementary envelope would be either  $Z = 2Y$  and the benefit of switching would be  $B = Z - Y = Y$ . Alternatively, if the complementary envelope were  $Z = Y/2$ , then the benefit of switching would be  $B = Z - Y = -Y/2$ . Since each of these outcomes is equally likely, the amount you stand to gain is greater than the

amount that you will lose, and the agent should switch. But, herein lies the paradox. For if the complementary envelope is treated as the randomly allocated envelope then the same reasoning applies.

It is merely coincidental that without considering the activities of the host and the consequent statistical events, the conclusions of many papers concur fortuitously with the findings presented in the subsequent chapters of the present paper.

In order for the agent of the two-envelope game to identify an envelope switching strategy, he or she must be informed of the host's activities, which are associated with the identification of the initial amount assigned to the first envelope.

It is demonstrated herein that once the agent has been informed about the distribution function, and upon sighting the content of the allocated envelope, he or she can derive and apply a suitable strategy to identify the expected benefit of switching. Consequently, the agent can decide whether to switch given the available information. Examples of different density functions are presented in Chapter 5.

## 1.4 Summary

The agent has to identify a game strategy that will realise the greatest expected benefit for each play of the game, either by retaining the randomly allocated envelope and its contents or by switching envelopes and accepting the content of the complementary envelope. The difference between the expected content of the complementary envelope and the observed content of the randomly allocated envelope is then the expected benefit of switching.

In the presence of the density function of the associated events for the observed amount, a benefit can be identified from switching the allocated envelope for the expected content of the complementary envelope. It is not the content of the complementary envelope, for any single trial, that dictates the strategy. It is the expected content of the complementary

envelope, for which the content of the allocated envelope is always noted to be  $Y = y$ , that generates the necessary switching strategy. The expected content of the complementary envelope is directly related to the initial amount assigned to the first envelope.

The agent, however, in revealing the content of the randomly assigned envelope, no longer needs to be concerned with the content of the complementary envelope. The sighted content is the means whereby the expected content of the complementary envelope is derived.

When the agent is not informed about the density function, he or she can use the information that each play of the game presents in order to refine an assumed prior distribution, which would be useful in subsequent plays of the game. Therefore, the approach adopted throughout this paper relies on the fact that a density function, whether derived or provided, is given.

In deriving his or her strategy, the agent considers the content of the randomly allocated envelope and the expected content of its complement. The expected difference between the contents of the former and the latter expresses the benefit to be realised should the agent decide to make this switch.

Thus, the agent assesses the available information associated with the content and allocation processes before adopting his or her chosen strategy.

Information on the activities associated with the host's generating of the game is essential for the agent to identify an appropriate and optimal playing strategy. Further, the agent is assumed to be risk-neutral and rational.

In summary, this research shows that, in the presence of information, it is possible for the agent to determine a switching strategy that can identify the expected benefit of switching envelopes for any observed amount in the randomly allocated envelope. Even after the play, the gained knowledge of the content of the envelopes has no bearing on

successive plays of the game if the agent is informed about the distribution function from which the host sampled the initial amount that was assigned to the first envelope.



# Chapter 2

## Literature Review

You know that envelope A contains a randomly determined sum of money, and envelope B contains, as a consequence of the flip of a fair coin, either half or twice that sum. And you know that one of these envelopes, randomly selected, has been given to you.

Suppose you are a risk-neutral money seeker who can choose between the money contained in the envelope that has been given to you and the money in the other envelope. Which do you prefer, or are you indifferent?

Suppose that you open your envelope. Now that you know exactly how much money is in it, which do you prefer, this money, or the money in the other envelope?

— Two Envelopes (Sobel 1994:p1)

The above-described dilemma posed by Sobel (1994:p1) is typical of the majority of papers that discuss the two-envelope problem, which as noted in Chapter 1, has engendered a great deal of research interest and debate. The challenge it has presented to many and the consequent explanations by authors have not necessarily produced a convincing argument to substantiate or refute the two-envelope paradox, however.

A story be told about an interesting paradox? But, perhaps even more interesting is the story that can be told of the many attempts to resolve an otherwise simple problem and thereby the paradox. That the paradox has attracted so much interest from such a diverse array of authors is a testament to the intrigue surrounding the problem and possibly academic prowess.

It is true that the amount of money in one envelope is twice that in the other envelope. But what is the implication of this statement that would seem to be quoted in nearly every publication. What is the true meaning of this statement and how can the other envelope always contain a greater expected amount of money regardless of which envelope you choose?

And it can be said that the observed amount is possibly the smaller or the larger of the content of the two envelopes. It can also be said that the content of the one envelope is twice the other, or half the other. These innocent statements only obfuscate what is the essential for the resolution of this problem.

What is necessary is that the reader recognises the provided information and then proceeds to understand the activities of the host of the game. In so doing, the agent will come to realise what additional information would make it essential for a winning strategy.

That the problem is ill-posed has been recognised by Loredó (2004:p2) and a re-formulation is necessary to facilitate the derivation of an appropriate switching strategy.

The alert reader will note the lack of consistency and the many attempts from diverse disciplines of those that have sought to settle the issue. We should heed the warning by Chihara (1995:p5) of the the subtle nature of the problem and exercise caution. Many are tempted to resolve the two-envelope problem through all manner of means when relying on increasingly complex arguments and may miss the point (Rawling 1994:p97).

## 2.1 Posing the problem

The variants, of which this version by Sobel (1994:p1) is but one, are essentially all the same. This is simply a mathematical problem that is not well-posed and over time many authors have recognised this and consequently re-formulated the problem and made it more tractable (Silver 1994:p19).

The presentation of the two-envelope problem rarely informs the agent about the content and allocation process that is adopted by the host. There is also some degree of confusion about the distinct difference between the amounts assigned to the first and second envelopes and what is contained in the allocated and complementary envelopes (Rawling 1994:p98).

Recent papers have referred to the density function from which the host selects the initial amount of money. If the agent is informed about this density function and understands the content and allocation process adopted by the host, then he or she can identify a strategy, provided that the content of the allocated envelope is sighted before deciding to switch envelopes or not.

Most authors in their exploration of the problem consider the amount of money in the allocated envelope and conclude that the amount in the complementary envelope is directly derivable from this sighted amount.

The real issue is being able to identify the amounts in the first and second envelopes that would have yielded the observed amount in the allocated envelope. It is the initial amount assigned to the first envelope and the consequent assigning of another amount to the second envelope followed by the random allocation of the indistinguishable envelopes that determines what the agent reveals in the allocated envelope.

### 2.1.1 The formulation

While, not being permitted to sight the content may be taken as information, it does render the decision to switch or not irrelevant and consequently the agent should be indifferent.

The realisation that in the absence of sighting the content of the allocated envelope there can be no benefit to switching the envelopes has been thoroughly reported by Loredó (2004:p2). But, it is the alternative situation, when the agent is afforded the opportunity of sighting the content of the allocated envelope that continues to frustrate and entertain (Broome 1995; Clark and Shackel 2000).

Once the host informs the agent of the content and allocation process, then the agent can derive a strategy. In the instance where the host withholds the information of the distribution function from which the initial amount would have been identified, then the agent can in the long run estimate and refine the prior distribution subsequent to each play of the game. As a result of this, it will be taken that the agent is informed of the distribution function. Strategies, in instances of the agent being informed, are readily identified by (Christensen and Utts 1992:p275), and (Brams and Kilgour 1995; Broome 1995; Bruss and Ruschendorf 2000; Castell and Batens 1994; Scott and Scott 1997).

It is not the expectation of the density function that is in question but the existence of the sub- $\sigma$ -algebra induced by the observed amount (Ash 1972:p254). Once this is accepted, then the expected benefit of switching is the conditional expectation of the sub- $\sigma$ -algebra induced by the observed amount. Any question of infinities can immediately be dismissed, and any concerns about the measures associated with contentious distribution functions are no longer valid (Nalebuff 1989:p176), (Rawling 1994:p98).

The concept of a sub- $\sigma$ -algebra ensures that the observed amount in the allocated envelope, also referred by Rawling (1994:p100), Jeffrey (1986:p8), and McGrew et al. (1997)

as a rigid designator, will allow the narrowing of the distance between opinions. Finally the consensus – albeit from different disciplines – will add credence to the overwhelming evidence refuting the existence of the paradox (Rawling 1994:p101).

### **2.1.2 The content**

What is assigned to the first and second envelope must be distinguished from what is in the allocated and complementary envelopes.

Most resort to describing the problem as consisting of two envelopes to which have been assigned sums of money. In the problem posed by Sobel (1994:p1) above, the amount assigned to the second envelope is either half or double the amount assigned to the first. This content and allocation process is also considered by Cargile (1992:p212), Nalebuff (1989), Bruss (1996), Rawling (1994:p97). By contrast Broome (1995) considers only that the amount assigned to the second envelope is double the amount assigned to the first envelope. While Brams and Kilgour (1995) assign an amount to the second envelope that is half the amount assigned to the first envelope. These different processes, of content and allocation, for the two-envelope problem yield distinctly different strategies, as will be demonstrated in Chapter 4.

The allocation strategy will be discussed in Chapter 3.

What process is responsible for observing an amount and then concluding that the other is half or double, is rarely considered. It is routinely taken that the relevant information is immediately before the agent and is all that is necessary to justify an appropriate strategy. Although the agent observes an amount of money, the amount observed is not what the host may have assigned to the first envelope. And this realisation makes it necessary for the agent to know the distribution function from which the initial amount was sampled (Loredo 2004:p5) . For example, in the instance of the “doubling-only” process, for the agent to observe a particular amount of money, it is possible that the

host may have sampled an initial amount that is exactly what the agent observes and the complementary envelope then contains double the observed amount. Or, the initial amount could have been half the amount observed and as a result of the “doubling-only” process the amount assigned to the second envelope is revealed as the content of the allocated envelope. This later alternative ensures that the complementary envelope contains half the the amount in the allocated envelope.

What matters is the content and allocation process and what the agent observes in the allocated envelope. With this in mind, the content of the complementary envelope at the time of play, is irrelevant to the agent.

The approach in this paper follows the discussion by Katz and Olin (2007:p914) in which they recognise that the amounts of money assigned to the envelopes, according to the content and allocation process, are relevant for the agent to derive an appropriate switching strategy.

We turn, now, to the fixed-sum formulation of the two envelope paradox, where the two mutually exclusive and jointly exhaustive possible states of the world that determine the consequences of your actions are characterised in terms of the actual contents of one of the envelopes, say A. (Katz and Olin 2007:p919)

Traditionally, in the version of the two-envelope problem where the agent is afforded the opportunity of sighting the content of the allocated envelope, the probability associated with the content being either the larger or the smaller is taken as the same. However, the probability of the allocated envelope containing the larger amount may not be the same as the probability of the envelope containing the smaller amount. Even considering this possibility detracts from the problem and in this paper it will be demonstrated that what one observes in the allocated envelope is a function of what was assigned to the first envelope and the issue is not whether the allocated envelope contains the smaller or the

larger amount. It is the process adopted by the host that dictates what the agent will observe, in the allocated envelope, that is now relevant. Is the envelope into which the initial amount assigned allocated to the agent before deciding what is to be assigned to the second envelope? In this instance, the “halving or doubling” method must be used, since if the “doubling-only” method were adopted, then the agent would know that the complementary envelope contained twice the content of the randomly allocated envelope, and it would be certain in all plays of the game that a switch would result in a guaranteed benefit.

It will be demonstrated in Chapter 3 and Chapter 4 that while it is true that the content of the complementary envelope is twice or half of the content of the allocated envelope, it is nonetheless meaningless and of no relevance to the resolution of the problem.

## 2.2 The method

It is also true that once the first and second envelopes are primed that each have equal likelihood of been allocated to the agent. But, as will be noted in the strategy chapter such probabilities are factored out in the expected benefit functions. The relevance of equal likelihood for the allocation of the two envelopes is questioned by (Albers et al. 2005:p92)

The contextualising of the dilemma in the form of a story necessitates that the reader must distill from the verbosity the essentials in order to formulate the problem. Smullyan (1992) tackles the problem by not embracing probability. A subtle wording of what can be gained or lost by switching would appear to be quite convincing in his particular argument. Because the alternatives are equally likely, probability should be considered (Chase 2002; Chihara 1995).

### 2.2.1 Reason and logic

Some authors such as Smullyan (1992) and Chihara (1995), pose the problem as a story and then proceed logically to solve it.

Chen (2007) and Rawling (1994:p102) argue that neither Proposition 1 nor Proposition 2, as proposed by Smullyan (1992:p190), are correct, and the very wording of the propositions is questioned by (Albers et al. 2005:p92).

It is acknowledged by many authors that no resolution to the paradox is possible with the available information. What is necessary is additional information, or the agent will need to make assumptions before attempting to resolve the problem. As such the problem is open to interpretation and the vague statement of the problem is probably the very reason that the literature is peppered with such diverse responses.

Some authors Cargile (1992), Smullyan (1992) and Chen (2007) tackle the problem using disciplines outside of the field of probability. Some papers (Sutton 2010:p10) of a philosophical nature acknowledge the need to resolve the problem by embracing a probabilistic argument. It will be noted that there is a convergence of argument and in many instances there is an agreement albeit from different disciplines. This confluence of ideas and argument are all lending credence to the demise of the apparent paradox.

### 2.2.2 Probability

The many different attempts to explain the legitimacy of the paradox stems from the fact that many authors are not well versed in the discipline of probability. The fundamentals essential to the probabilistic approach would seem to be taken for granted and in many instances the mere referencing to Bayesian or conditional probabilities would seem to add an air of legitimacy to the authors intent. The narrative that is used to deliver the problem is not very clear and care, as advocated by Vazquez (2012:p2), needs to be exercised by



the reader who is tempted to claim that he or she can resolve the two-envelope problem and thereby explain the paradox. In many instances attempts to resolve the paradox only reinforces and perpetuates the myth.

According to Broome (1995); Jackson et al. (1994), the distribution from which the host samples the initial amount is essential for the agent to derive an appropriate switching strategy. The distribution function that represents the initial monetary amount from which the host will sample, can be either discrete or continuous; this however is never mentioned in the usual statement of the problem. As shown in Chapter 4, this information is essential for the agent to derive an appropriate density function. For any successful strategy, the density function must be known by the agent to support the identification of an appropriate switching strategy (Jackson et al. 1994).

While many authors recognise the conditional probability (Bayesian) is necessary to derive an appropriate switching strategy, they unfortunately fail to recognise that the expected benefit is conditional on the sub- $\sigma$ -algebra induced by the amount observed in the allocated envelope (Ash 1972:p249). This necessitates that the agent be informed of the distribution function from which the host sampled the initial amount that was assigned to the first envelope. The resultant event space immediately lends itself readily to a solution that overcomes many of the concerns that have been expressed, as a means to justify or dismiss the paradox (Ok 2007:p12). Concerns, about infinite amounts of money that are labelled as absurd and devoid of all reality and therefore dismissed are no longer an issue when the event space is recognised as a sub- $\sigma$ -algebra. Whether the amounts of money in the financial system are infinite - the FED can always print more - or not, is besides the point (Rawling 1994).

The expected benefit is succinctly expressed as the conditional expectation for the sub- $\sigma$ -algebra induced by the observed amount, as discussed by Spanos (2013:p3) and Ash (1972:p249).

## 2.3 Switching strategies

In instances where the amounts are bounded, strategies can be derived without much effort, as will be detailed in Chapter 4.

It can be said that whether a distribution function has a finite mean or not is irrelevant for the identification of an appropriate switching strategy. Broome (1995) correctly states the problem that the agent is confronted with as one where he or she must consider the expected content of the complementary envelope once the content of the allocated envelope is revealed.

In some instances, the density function may support the notion that switching is always beneficial. In no way is this paradoxical; it is rather a consequence of the density function. Any density functions with a finite mean will not support the paradox. Broome (1995:p8) states that a switching strategy can be identified that will yield the greatest expected benefit for any observed amount from the allocated envelope. While distribution functions possessing no finite mean will always support the paradox, they render sighting the content of the allocated envelope redundant and support the notion that switching is always beneficial.

If the amounts assigned to the envelopes by the host are bounded, then knowledge of this by the agent is necessary. In such an instance, the paradox would not exist since it would not be beneficial for the agent to switch if the content of the allocated envelope were found to be greater than half the upper bound. Should the observed amount be less than twice the lower bound, however, then it would be advantageous to switch.

It is possible for the agent, upon sighting the content and observing the amount  $Y = y$ , to determine the conditional expectation of the benefit when switching envelopes. It will be realised that the conditional expectation is a function of the sub- $\sigma$ -algebra induced by  $Y$ . This realisation immediately eliminates some of the absurdities that have been espoused

as the cause of the paradox. Absurdities such as the infinite sets, are immediately reduced to finite sets, as a result of the induced sub- $\sigma$ -algebra (Spanos 2013).

It has been proposed by McDonnell and Abbott (2009:p3,p12), Bruss and Ruschendorf (2000:p7), and Gill (2011:p14) that a strategy could be based on “Cover’s switching function”. This will be less than ideal in some instances and an exact switching strategy can be derived from the provided distribution function. In the absence of a provided distribution function, through an adaptive strategy, the agent can assume a distribution whose parameters are refined upon the completion of each game.

It is also worth mentioning, as proved by Broome (1995:p10), that no distribution function possessing a finite mean can be “paradoxical”. In such instances there will be an interval over which it would be beneficial to switch envelopes and alternative intervals where switching would not be beneficial. Similarly, in the presence of bounds, it will not always be beneficial to switch.

### **2.3.1 The paradox**

Most authors attempt to resolve the paradox by comparing the content of the two envelopes. For whatever the content of the envelopes, the complementary envelope will always be half or twice the content of the allocated envelope. This justifies the switching. In this paper the content of the allocated envelope is treated as a constant (rigid designator) and the possibility of the content of the allocated envelope taking on the larger or the smaller amount is never in doubt. As such this aligns with the philosophical approach proposed by authors McGrew et al. (1997), Rawling (1994), and Priest and Restall (2007), who claim to have settled the issue of the paradox. In most of these instances they dismiss the paradox on the grounds that the content of the allocated envelope must be treated as a rigid designator (Chase 2002; Chen 2007; Sutton 2010). These arguments call on the disciplines of logic and are in this paper substantiated by the appropriate

application of probabilistic argument.

Sutton (2010) recognises the need for a probabilistic approach and that the inevitable resolution of the two-envelope problem will be understood by philosophers.

Broome (1995) approaches the paradox by considering that the smaller monetary amount is essential and then proceeds to derive a function that expresses the expectation as a result of switching. There are two distributions presented by Broome (1995) to support the strategy of always switching. Broome (1995) acknowledges that these distributions, although valid, possess no finite mean, and demonstrates that no distribution with a finite mean (even an infinite distribution) can be paradoxical.

## 2.4 Summary

As will be noted from the aforementioned there are any number of attempts to resolve the two-envelope problem from a diverse array of disciplines. Many practitioners have relied on an extensive array of simple and complex arguments to justify their conclusions. And in most instances there is a reliance on a mathematical formulation of the problem. It is this formulation that I believe is the cause of the paradox. As will be seen, provided that there is information, then a strategy can be identified that will yield a greatest expected benefit for any play of the game.

# Chapter 3

## Posing the two-envelope problem

What is essential is invisible to the eye.

— The Little Prince (de Saint-Exupéry 1945:p70)

The game of switching envelopes, which lends its name to the two-envelope paradox, is based on the problem that confronts an agent who must identify a strategy to maximise the expected benefit by either retaining the allocated envelope and its contents or choosing the complementary envelope and its contents. The host of the game is responsible for assigning the amounts of money, according to a specified content and allocation process, to each of the two envelopes. The money assigned to the first envelope is a randomly identified amount that is distributed according to some probability density function. The second amount is dependent upon the amount assigned to the first envelope; it could be either half or double that amount depending on the “content” component of the content and allocation process.

The random allocation of the two indistinguishable envelopes will see the agent receiving one of them. Neither the agent nor the host at the time of allocation can discern the contents and neither would they, until the contents are revealed, be in a position to comment on the appropriate switching strategy.

We divide the two-envelope game into two stages: the activities of the host (stage 1) and the strategy of the agent (stage 2). The host of the game conducts three experiments in stage 1. The first experiment is the random sampling from an arbitrary distribution function, the second experiment is the tossing of an unbiased coin, and the third and final experiment is the tossing of the same coin again. In stage 2, the agent identifies the envelope-switching strategy that yields the greatest expected benefit, which is determined by considering the expected difference between the content of the complementary envelope and the content of the allocated envelope.

Bruss and Ruschendorf (2000) and Broome (1995) are among the authors to state that the content (the smaller amount) of the first envelope is identified by the host and that a subsequent amount that is twice the initial amount is assigned to a second envelope. On the other hand, Sobel (1994:p69) refers specifically to halving or doubling the initial amount depending on the outcome of a random event. In yet other instances, no consideration is paid to the content and allocation process. However, this paper demonstrates that distinguishing between the two processes is important as they may yield distinctly different strategies. Where sighting the content of the randomly allocated envelope is prohibited, the content allocation process is irrelevant (Bruss 1996:p117).

### 3.1 The host's activities

The concept of a probability space  $(\Omega, \mathcal{F}, P)$  is introduced, where  $\Omega$  is the sample space,  $\mathcal{F}$  the event space, and  $P$  the probability measure on  $\mathcal{F}$ . All subsets of  $\Omega$  are events, and  $\mathcal{F}$  consists of all subsets of  $\Omega$ .

Experiment 1, performed by the host, involves identifying the initial amount of money that is assigned to the first envelope. The initial amount is drawn from a cumulative distribution function, which is either given or derived from an arbitrary probability density

function.

The outcome of Experiment 2, performed by the host, is used to decide whether to halve or double the initial amount and assign the resultant amount to the second envelope.

The outcome of Experiment 3, performed by the host, is used to decide whether to allocate the first or the second envelope to the agent. The remaining envelope will be referred to as the complementary envelope.

The grand experiment, for consideration by the agent, and the associated product probability space is a function of these three independent host-conducted experiments. Associated with each of them is an appropriate probability space that would need careful consideration by the agent in order for him or her to identify an appropriate strategy.

**Experiment 1.**  $X_1$  is a random variable on  $(\Omega_1; \mathcal{F}_1; P_1)$ .  $X_1$  represents the initial amount of money identified by the host and assigned to the first of the two envelopes.

The sample space associated with the drawing of an initial random amount  $x_1$  represented by the random variable  $X_1$ , where  $X_1 \sim f_{X_1}(x_1)$ , is defined as:

$$\Omega_1 = \{\omega_1: x_l < X_1(\omega_1) \leq x_u\}.$$

This initial amount  $X_1(\omega_1) = x_1$  is assigned to the first of the two envelopes. Unless otherwise specified, the notation  $X_1(\omega) = x_1$  is hereafter represented by  $X_1 = x_1$ . The sample space for drawing the initial amount is determined by the host. We can see it as the values between  $x_l$  and  $x_u$ .

**Experiment 2.**  $X_2$  is a random variable on  $(\Omega_2; \mathcal{F}_2; P_2)$ .  $X_2$  represents the outcome of the toss of an unbiased coin. The sample space is

$$\Omega_2 = \{0; 1\}$$

$X_2(\omega_2) = \omega_2$  and  $P(\omega_2) = 1/2 \quad \forall \omega_2 \in \{0; 1\}$ .

$$\omega_2 = \begin{cases} 0, & \text{halve the initial amount;} \\ 1, & \text{double the initial amount.} \end{cases}$$

**Experiment 3.**  $X_3$  is a random variable on  $(\Omega_3; \mathcal{F}_3; P_3)$ .  $X_3$  also represents the outcome of the toss of an unbiased coin. The sample space is again

$$\Omega_3 = \{0; 1\}$$

$X_3(\omega_3) = \omega_3$  and  $P(\omega_3) = 1/2 \quad \forall \omega_3 \in \{0; 1\}$ .

$$\omega_3 = \begin{cases} 0, & \text{allocate the first envelope to the agent;} \\ 1, & \text{allocate the second envelope to the agent.} \end{cases}$$

## 3.2 The agent

The probability space that the agent would need to consider is

$$(\Omega; \mathcal{F}; P) = (\Omega_1 \times \Omega_2 \times \Omega_3; \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3; P_1 \times P_2 \times P_3)$$

and the sample space for the agent (product measure space of the host) is expressed as

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3.$$

Since the experiments are conducted independently, it can be said that  $P = P_1 \times P_2 \times P_3$  and the proof of the existence of such a  $P$  on  $\mathcal{F}$  is an application of Carathéodory's extension theorem (Athreya and Lahiri 2010:p24).

For any outcome  $\{c\} = \{x_1; \omega_2; \omega_3\} \in \mathcal{F}$ , since the outcomes (events)  $x_1$ ,  $\omega_2$ , and  $\omega_3$



originate from each of the independent experiments, then

$$\begin{aligned}
P(\{c\}) &= P(\{x_1; \omega_2; \omega_3\}) \\
&= P(X_1 = x_1; X_2 = \omega_2; X_3 = \omega_3) \\
&= P(X_1 = x_1)P(X_2 = \omega_2)P(X_3 = \omega_3).
\end{aligned} \tag{3.1}$$

### 3.3 Envelope contents

Knowledge of the process that the host adopted for the priming of the envelopes is crucial for the agent to determine an appropriate switching strategy. Most formulations of the problem centre on the “doubling-only” or “halving or doubling” processes. A “halving-only” process is introduced as the analog of the “doubling-only” process and it will assist with the derivation of the different switching strategies.

Having randomly identified an initial amount  $X_1 = x_1$ , the host assigns this to the first of the two indistinguishable envelopes. To the second envelope, the host will assign an amount  $X'_1$  that is dependent on the outcome of the toss of an unbiased coin. The transformation associated with the halving or doubling of the initial amount is as follows:

$$X'_1 = \left( \frac{1}{2}(1 - X_2) + 2X_2 \right) X_1 \tag{3.2}$$

where

$$X_2 = \begin{cases} 0, & \text{halve the initial amount;} \\ 1, & \text{double the initial amount.} \end{cases}$$

Thus, the probability of any outcome  $\{c\} = \{x_1; \omega_2; \omega_3\}$  associated with the “halving or doubling” process after applying 3.1 is

$$P(\{c\}) = \frac{1}{4}P(X_1 = x_1). \quad (3.3)$$

The “halving or doubling” process is a combination of the “halving-only” and the “doubling-only” processes. Each of these processes can be realised by setting the relevant parameter  $\omega_2$ .

### 3.3.1 Doubling-only

For the “doubling-only” process, the host only doubles the content of the first envelope and assigns this amount to the second envelope. Formally,  $X'_1 = 2X_1$ .

This can be realised by setting  $X_2(\omega_2) = \omega_2 = 1$ .  $\Omega_2 = \{1\}$  and  $P(\omega_2) = 1$ . Therefore, noting that  $\{c\} = \{x_1; 1; \omega_3\}$ , after applying (3.1)

$$P(\{c\}) = \frac{1}{2}P(X_1 = x_1). \quad (3.4)$$

### 3.3.2 Halving-only

For the “halving-only” process, the host only halves the content of the first envelope and assigns this amount to the second envelope. Formally,  $X'_1 = X_1/2$ .

This can be realised by setting  $X_2(\omega_2) = \omega_2 = 0$ .  $\Omega_2 = \{0\}$  and  $P(\omega_2) = 1$ . Therefore, noting that  $\{c\} = (x_1; 0; \omega_3)$ , after applying (3.1)

$$P(\{c\}) = \frac{1}{2}P(X_1 = x_1). \quad (3.5)$$

### 3.4 The envelope allocation

After priming the envelopes, the host randomly selects one of the indistinguishable envelopes and allocates this envelope to the agent. The agent's envelope and its associated contents  $y$  are represented by the random variable  $Y$ , while the complementary envelope and its associated contents  $z$  are represented by the random variable  $Z$ .

The transformation process associated with the allocation of the envelopes and their respective contents can be expressed as follows:

$$\begin{aligned} \mathbf{C} &= \begin{bmatrix} Y & Z \end{bmatrix} \\ &= \begin{bmatrix} X_1 & X'_1 \end{bmatrix} \begin{bmatrix} \mathbf{A} \end{bmatrix} \\ &= X_1 \begin{bmatrix} 1 & \frac{1}{2}(1 - X_2) + 2X_2 \end{bmatrix} \begin{bmatrix} \mathbf{A} \end{bmatrix} \end{aligned} \quad (3.6)$$

where

$$\mathbf{A} = \begin{bmatrix} 1 - X_3 & X_3 \\ X_3 & 1 - X_3 \end{bmatrix},$$

$$X_2 = \begin{cases} 0, & \text{halve the initial amount;} \\ 1, & \text{double the initial amount,} \end{cases}$$

and

$$X_3 = \begin{cases} 0, & \text{allocate the first envelope to the agent;} \\ 1, & \text{allocate the second envelope to the agent.} \end{cases}$$

### 3.5 The benefit of switching envelopes

Since the objective of the agent is to identify an envelope-switching strategy that yields the greatest expected benefit for any observed  $Y = y$ , it is convenient to expand (3.6) and accordingly derive the benefit function in the event of switching the randomly allocated envelope for the complementary envelope.

The content of the agent's envelope is

$$Y = X_1 \begin{bmatrix} 1 & \frac{1}{2}(1 - X_2) + 2X_2 \end{bmatrix} \begin{bmatrix} 1 - X_3 \\ X_3 \end{bmatrix}. \quad (3.7)$$

The content of the complementary envelope is

$$Z = X_1 \begin{bmatrix} 1 & \frac{1}{2}(1 - X_2) + 2X_2 \end{bmatrix} \begin{bmatrix} X_3 \\ 1 - X_3 \end{bmatrix}. \quad (3.8)$$

The benefit of switching envelopes is

$$\begin{aligned}
B &= Z - Y \\
&= \mathbf{C} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} X_1 & X'_1 \end{bmatrix} \begin{bmatrix} 2X_3 - 1 \\ 1 - 2X_3 \end{bmatrix} \\
&= \frac{1}{2}(2X_3 - 1)(1 - 3X_2)X_1.
\end{aligned} \tag{3.9}$$

It will be noted that this expression makes no direct reference to the content of the complementary envelope. The benefit is therefore entirely dependent on the initial amount identified by the host and the halving or doubling of this initial amount to determine the content of the second envelope. Finally, the random allocation of the envelopes, determines what is revealed in the allocated envelope.

### 3.6 Summary

The host, having performed three experiments, assigns the initial amount to the first of the two envelopes and then assigned an amount to the second envelope, which is a function of the first amount. One of the primed envelopes is then randomly allocated to the agent.

The agent may receive the envelope containing  $X_1 = x_1$  and refer to the content as “Y”; alternatively, he or she may receive the envelope containing the amount associated with  $X'_1$  and similarly refer to this as “Y”.

In all instances, the allocation process will commence with the identification of an initial amount  $x_1$ , represented by the random variable  $X_1$ , which can be assumed to originate

from a density function  $f_{X_1}(x_1)$ . If the agent is not informed of the distribution function  $F_{X_1}(x_1)$  from which the host sampled the initial amount, then the agent can assume or estimate the density function  $f_{X_1}(x_1)$ . Suffice to say that the agent will be informed and can in the presence of the density function  $f_{X_1}(x_1)$ , together with the content and allocation process derive an appropriate switching strategy.

## Chapter 4

# The expected benefit of information

Words are the source of misunderstandings.

— The Little Prince (de Saint-Exupéry 1945:p65)

A switching strategy that optimises the expected outcome for any observed amount is possible if the agent is provided with information about the distribution function from which the initial amount was identified. This provision of information implies that the agent is informed about the content and allocation process as well as the distribution function from which the host randomly selected the initial amount. Not only is the distribution function essential for the agent to identify a strategy, but the content and allocation process is also crucial, as will be demonstrated in the strategy associated with the halving or doubling process. With this information, the agent can derive a game strategy that will enable him or her, for any observed amount, to identify the expected benefit and in so doing possess the information that will facilitate the decision to retain the allocated envelope and its contents or to switch to the complementary envelope and its contents.

## 4.1 Essential information for the agent

If the agent is not informed about the distribution function from which the original  $X_1 = x_1$  was selected, he or she may assume the existence of a prior density function. In addition, consequent to each play of the game, the agent may revise the guessed prior density function for the subsequent play of the game. However, irrespective of whether the agent has derived such a density function or not, the density function under consideration will be referred to as  $f_{X_1}(x_1)$ .

Without loss of generality, any reference to distribution will imply a cumulative distribution function whether associated with a discrete – or a continuous random variable. Similarly, there will be no notational distinction between a probability mass function associated with a discrete random variable and the alternative probability density function associated with a continuous random variable. The terms “discrete” and “continuous” are used to indicate the nature of the density function and similarly the distribution function. Since the objective of the game is to identify a strategy that will yield the greatest expected benefit, it is necessary to assume that all plays of the game are made by a risk-neutral and rational agent.

The agent will need to consider the content and allocation process adopted by the host and nature of the distribution function from which the host will sample the initial amount that will be assigned to the first envelope. The only information available to the host and the agent once the envelopes are allocated is the distribution from which the host sampled, the content and allocation process. The fact that the content of one envelope is double the content of the other, while true, is of no significance in determining the appropriate switching strategy. But, if one were informed that the agent had adopted the “doubling-only” process then it would be correct to say that the content of the second envelope is primed by doubling the amount assigned to the first envelope.



The agent will need to identify — for the observed amount  $y$  revealed in the randomly allocated envelope represented by the random variable  $Y$  — the expected benefit to be had by exercising the option of switching envelopes and therefore obtaining the expected content of the complementary envelope  $Z$ .

The benefit is represented by the random variable  $B$  and is derived from the difference  $B = Z - Y$  between the amount (not known) in the complementary envelope, represented by the random variable  $Z$ , and the content (observed) of the randomly allocated envelope, represented by the random variable  $Y$ .

## 4.2 The events associated with the sighted amount

For the agent to observe an amount  $Y = y$  the host would have had to have sampled an initial amount that may have been any one of the members of the set  $x_1 \in \{y/2; y; 2y\}$ . It is the informed agent that can derive an appropriate switching strategy. This is immediately recognisable when the agent considers the sub- $\sigma$ -algebra of relevant events induced by the observed amount  $Y = y$  (Ash 1972). The conditional expected benefit of switching the allocated envelope for the complementary envelope is then easily derived when the appropriate event space is considered (Spanos 2013).

$$x_1 \in \begin{cases} \left\{ \frac{y}{2}; y \right\}, & \text{for “doubling-only”;} \\ \left\{ y; 2y \right\}, & \text{for “halving-only”;} \\ \left\{ \frac{y}{2}; y; 2y \right\}, & \text{for “halving or doubling”}. \end{cases}$$

Suffice it to say, the contents of the envelopes are fixed and therefore the agent will reveal the content of the randomly allocated envelope to be  $Y = y$ . The agent will not know if this is the smaller amount and that the complementary envelope contains  $Z = 2y$  or the larger amount and that the complementary envelope then contains  $Z = y/2$ . While

this statement may appear to relevant it will be revealed that this is nothing more than an interesting comment with no relevance to the agent when deriving the appropriate switching strategy. What is now relevant is the original amount identified by the host and how the content and allocation process impacts on this initial amount that ultimately determines what will be revealed in the envelope that is allocated to the agent.

The following examples serve to explain the necessity of the agent, when observing  $Y = y$  in the randomly allocated envelope, to consider that the host — having adopted a halving or doubling process — may have selected an initial amount  $x_1 \in \{y/2; y; 2y\}$ . Each of these possibilities  $x_1 \in \{y/2; y; 2y\}$  is essential for the agent to identify a suitable switching strategy to realise the (greatest) expected benefit for any observed  $Y = y$  in the randomly allocated envelope.

Consideration will be given to the “doubling-only” strategy and, for the sake of expediency, a “halving-only” strategy will be introduced. By combining these two strategies a “halving or doubling” strategy will be derived. The “halving-only” strategy, as the complement of the “doubling-only” strategy, will be recognised as a distinct yet significant component of the “halving or doubling” strategy. To facilitate the derivation of the respective strategies for both discrete and continuous variables the concept of a sub- $\sigma$ -algebra, induced by  $Y$ , will be introduced for each of the strategies.

What is relevant to the agent, upon sighting the content of the allocated envelope, is not the event space  $\mathcal{F}$  but the event space  $\mathcal{F}'$  induced by the observed amount  $Y$  (Ash 1972:p250). With this information the agent is able to determine the expected benefit of switching for any observed  $Y = y$  (Ok 2007:p12).

#### 4.2.1 A “doubling-only” strategy

The agent observes  $Y = y$  and considers the effects of the host having selected  $x_1 \in \{y/2; y\}$ . It is now essential for the agent to determine the values of  $X_3 = \omega_3$  for which

the allocated envelope would contain  $Y = y$ , when considering each of the initial assigned amounts  $x_1 \in \{y/2; y\}$ .  $X_2 = \omega_2 = 1$ .

**Case 1.** *The host assigns the amount  $X_1 = y/2$  to the first envelope . The host doubles the initial amount and assigns this amount  $X'_1 = 2X_1 = y$  to the second envelope, and then allocates this second envelope to the agent, who will observe the amount  $Y = y$ .*

By substituting  $X_1 = y/2$ ,  $Y = y$ . and  $\omega_2 = 1$  into (3.7) and solving for  $\omega_3$ , we find

$$y = ((1 - \omega_3) + 2\omega_3)\frac{y}{2}.$$

The only possible solution to this equation is when  $\omega_3 = 1$  (allocating the second envelope to the agent). The event  $\{c\} = \{y/2; 1; 1\}$  is thus realised; consequently, by substitution into (3.8),  $Z = y/2$  and therefore

$$b(\{c\}) = b(\{y/2; 1; 1\}) = -y/2$$

and

$$P(\{c\}) = P(\{y/2; 1; 1\}) = \frac{1}{2}P(X_1 = y/2).$$

**Case 2.** *The host assigns the amount  $X_1 = y$  to the first envelope .The host doubles the initial amount and assigns this amount  $X'_1 = 2X_1 = 2y$  to the second envelope. If the first envelope is allocated, observe the amount  $Y = y$ .*

By substituting  $X_1 = y$ ,  $Y = y$ , and  $\omega_2 = 1$  into (3.7) and solving for  $\omega_3$ , we find

$$y = ((1 - \omega_3) + 2\omega_3)y.$$

The only possible solution to this equation is when  $\omega_3 = 0$  (allocating the first envelope to the agent). The event  $\{c\} = \{y; 1; 0\}$  is thus realised; consequently, by substitution into (3.8),  $Z = 2y$  and therefore

$$b(\{c\}) = b(\{y; 1; 0\}) = y$$

and

$$P(\{c\}) = P(\{y; 1; 0\}) = \frac{1}{2}P(X_1 = y).$$

Consequently, the set of events for consideration are

$$C_y = \{\{x_1; \omega_2; \omega_3\} : \{y/2; 1; 1\}; \{y; 1; 0\}\}. \quad (4.1)$$

#### 4.2.2 A “halving-only” strategy

The agent observes  $Y = y$  and considers the effects of the host having selected  $x_1 \in \{y; 2y\}$ . It is now essential for the agent to determine the values of  $X_3 = \omega_3$  for which the allocated envelope would contain  $Y = y$ , when considering each of the initial assigned amounts  $x_1 \in \{y; 2y\}$ .  $X_2 = \omega_2 = 0$ .

**Case 3.** *The host assigns the amount  $X_1 = y$  to the first envelope. The host halves the initial amount and assigns this amount  $X'_1 = y/2$  to the second envelope. If the first envelope is allocated, then the agent will observe the amount  $Y = y$ .*

By substituting  $X_1 = y$ ,  $Y = y$ , and  $\omega_2 = 0$  into (3.7) and solving for  $\omega_3$ , we find

$$y = (1 - \frac{1}{2}\omega_3)y.$$

The only possible solution to this equation is when  $\omega_3 = 0$  (allocating the first envelope to the agent). The event  $\{c\} = \{y; 0; 0\}$  is thus realised; consequently, by substitution into (3.8),  $Z = y/2$  and therefore

$$b(\{c\}) = b(\{y; 0; 0\}) = -y/2$$

and

$$P(\{c\}) = P(\{y; 0; 0\}) = \frac{1}{2}P(X_1 = y).$$

**Case 4.** *The host assigns the amount  $X_1 = 2y$  to the first envelope. The host halves the initial amount and assigns this amount  $X'_1 = y$  to the second envelope, and then allocates this second envelope to the agent, then the agent will observe the amount  $Y = y$ .*

By substituting  $X_1 = 2y$ ,  $Y = y$ . and  $\omega_2 = 0$  into (3.7) and solving for  $\omega_3$ , we find

$$y = (1 - \frac{1}{2}\omega_3)2y.$$

The only possible solution to this equation is when  $\omega_3 = 1$  (allocating the second envelope to the agent). The event  $\{c\} = \{2y; 0; 1\}$  is thus realised; consequently, by substitution into (3.8),  $Z = 2y$  and therefore

$$b(\{c\}) = b(\{2y; 0; 1\}) = y$$

and

$$P(\{c\}) = P(\{2y; 0; 1\}) = \frac{1}{2}P(X_1 = 2y).$$

Consequently, the set of events for consideration are

$$C'_y = \{\{x_1; \omega_2; \omega_3\} : \{y; 0; 0\}; \{2y; 0; 1\}\}. \quad (4.2)$$

### 4.2.3 A “halving or doubling” strategy

This strategy, it will be realised, is a combination of the “halving-only” and the “doubling-only” strategies.

The agent observes  $Y = y$  and considers the effects of the host having selected  $x_1 \in \{y/2; y; 2y\}$ . It is now essential for the agent to identify the events that would describe the host’s activities, namely determine the values of  $\omega_2$  and  $\omega_3$  for which the allocated envelope would contain  $Y = y$ , when he or she considers each of the initial assigned amounts  $x_1 \in \{y/2; y; 2y\}$ .

The combined events associated with the “halving-only” and “doubling-only” process are necessary for the “halving or doubling” process.

$$\begin{aligned}\mathcal{C}_y &= C_y \cup C'_y \\ &= \{\{x_1; \omega_2; \omega_3\} : \{y/2; 1; 1\}; \{y; 0; 0\}; \{y; 1; 0\}; \{2y; 0; 1\}\}.\end{aligned}\tag{4.3}$$

### 4.2.4 Expected benefit associated with discrete variables

The events of relevance to the agent for the identification of a strategy that will realise the expected benefit for any observed amount  $Y = y$  are the members (events) of  $\mathcal{C}_y$ ,  $C'_y$ , or of  $C_y$ , depending on which content and allocation has been adopted by the host.

For the “doubling-only” strategy

$$\begin{aligned}
E(B|Y = y) &= E(B|C_y) \\
&= \frac{\sum_{\{c\} \in C_y} b(\{c\})P(\{c\})}{\sum_{\{c\} \in C_y} P(\{c\})} \\
&= \frac{-\frac{y}{2}P(X_1 = \frac{y}{2}) + yP(X_1 = y)}{P(X_1 = \frac{y}{2}) + P(X_1 = y)}.
\end{aligned} \tag{4.4}$$

For the “halving-only” strategy

$$\begin{aligned}
E(B|Y = y) &= E(B|C'_y) \\
&= \frac{\sum_{\{c\} \in C'_y} b(\{c\})P(\{c\})}{\sum_{\{c\} \in C'_y} P(\{c\})} \\
&= \frac{-\frac{y}{2}P(X_1 = y) + yP(X_1 = 2y)}{P(X_1 = y) + P(X_1 = 2y)}.
\end{aligned} \tag{4.5}$$

For the “halving or doubling” strategy, from (4.4) and (4.5)

$$\begin{aligned}
E(B|Y = y) &= E(B|\mathcal{C}_y) \\
&= \frac{\sum_{\{c\} \in \mathcal{C}_y} b(\{c\})P(\{c\})}{\sum_{\{c\} \in \mathcal{C}_y} P(\{c\})} \\
&= \frac{\sum_{\{c\} \in C_y \cup C'_y} b(\{c\})P(\{c\})}{\sum_{\{c\} \in C_y \cup C'_y} P(\{c\})} \\
&= \frac{\sum_{\{c\} \in C_y} b(\{c\})P(\{c\}) + \sum_{\{c\} \in C'_y} b(\{c\})P(\{c\})}{\sum_{\{c\} \in C_y} P(\{c\}) + \sum_{\{c\} \in C'_y} P(\{c\})} \\
&= \frac{-\frac{y}{2}P(X_1 = \frac{y}{2}) + \frac{y}{2}P(X_1 = y) + yP(X_1 = 2y)}{P(X_1 = \frac{y}{2}) + 2P(X_1 = y) + P(X_1 = 2y)}. \tag{4.6}
\end{aligned}$$

### 4.3 Continuous random variables

All discussions have centred on the agent observing an amount that is represented by a discrete random variable. It is possible that the observed amount can be represented by a continuous random variable. The probability of any observed amount represented by a continuous random variable, as is known, is zero (Ok 2007:p12). However, if the observed amount is centred on an interval, the length of which tends to zero, then it is possible to derive a function for the continuous random variable analogous to that for the discrete random variable. This variant of the problem will permit the derivation of a strategy associated with an amount that is represented by a continuous random variable.

Probabilities associated with continuous variables require a little more care in order to facilitate the derivation of appropriate switching strategies. Probabilities associated with intervals are easily dealt with for continuous variables. This feature will be exploited by considering intervals of lengths, that are proportionately  $\varepsilon$ ,  $2\varepsilon$ , and  $4\varepsilon$  in length.

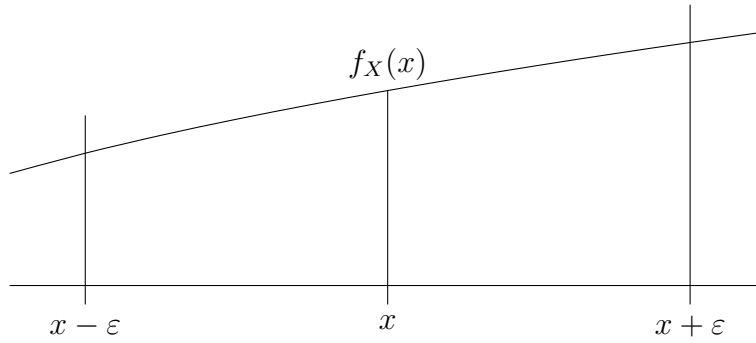


For any chosen interval, the limit as  $\varepsilon$  tends to zero will be exploited. As noted by (Ash 1972:p241), (Mood et al. 1974:p61), (Brams and Kilgour 1995:p30), and (Broome 1995:p10), the approximation does yield the appropriate switching function.

Where  $f_{X_1}(x)$  is continuous, the following approach is adopted. For the agent to observe an amount in the allocated envelope, from the interval  $y - \varepsilon < Y \leq y + \varepsilon$ , the agent would have to consider the activities of the host in drawing, depending on the content and allocation process, an initial amount from the set of intervals:

### 4.3.1 Probability

$P(X = x)$  is not true for a continuous probability density function  $f_X(x)$ . However, it can be approximated as described by Ash (1972:241–261) and Mood et al. (1974:61).



$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{F_X(x + \varepsilon) - F_X(x - \varepsilon)}{2\varepsilon} \end{aligned}$$

Hence,

$$\begin{aligned}
f_X(x)2\varepsilon &\approx F_X(x+\varepsilon) - F_X(x-\varepsilon) = P(x-\varepsilon < X \leq x+\varepsilon) \\
&= P(X \leq x+\varepsilon) - P(X \leq x-\varepsilon) \\
&= \int_{x-\varepsilon}^{x+\varepsilon} f_X(x_1) dx_1
\end{aligned}$$

### 4.3.2 Expectation

The  $\sigma$ -algebra induced by the observed amount is very convenient. For the random object  $Y : (\Omega; \mathcal{F}) \longrightarrow (\Omega'; \mathcal{F}')$ , then the  $\sigma$ -algebra induced by  $Y$  is given by  $\mathcal{F}(Y) = Y^{-1}(\mathcal{F}')$  (Ash 1972:p250). And, the conditional expectation  $E(B|Y = y)$  is well defined for any number  $Y = y$  even though  $Y = y$  may be a probability zero event (Ok 2007:p12).

It would be beneficial to derive an appropriate definition for the expectation of continuous variable  $X$  over a stated interval  $y - \varepsilon < X \leq y + \varepsilon$  (Rényi 1970:265). Then

$$\begin{aligned}
E(X = y) &\approx \lim_{\varepsilon \rightarrow 0} E(X | y - \varepsilon < X \leq y + \varepsilon) \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\int_{y-\varepsilon}^{y+\varepsilon} x f_X(x) dx}{\int_{y-\varepsilon}^{y+\varepsilon} f_X(x) dx} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{y f_X(y) 2\varepsilon}{f_X(y) 2\varepsilon} \\
&= y
\end{aligned}$$

### 4.3.3 A “doubling-only” strategy

For the “doubling-only” process

$$C_y = \bigcup_{\omega_3 \in \{0;1\}} \{ \{x_1, 1, \omega_3\} : 2^{-\omega_3}(y - \varepsilon) < x_1 \leq 2^{-\omega_3}(y + \varepsilon) \}.$$

Noting, from (3.9), that

$$b(\{c\}) = b(\{x_1; 1; \omega_3\}) = \begin{cases} x_1, & \text{for } \omega_3 = 0; \\ -x_1, & \text{for } \omega_3 = 1, \end{cases}$$

and following from (3.1) then

$$\begin{aligned} E(B|Y = y) &\approx \lim_{\varepsilon \rightarrow 0} E(B|y - \varepsilon < Y \leq y + \varepsilon) \\ &= \lim_{\varepsilon \rightarrow 0} E(B|C_y) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\int_{C_y} b(\{c\}) f_{X_1}(x_1) dx_1}{\int_{C_y} f_{X_1}(x_1) dx_1} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{-\int_{\frac{y}{2} - \frac{\varepsilon}{2}}^{\frac{y}{2} + \frac{\varepsilon}{2}} x_1 f_{X_1}(x_1) dx_1 + \int_{y - \varepsilon}^{y + \varepsilon} x_1 f_{X_1}(x_1) dx_1}{\int_{\frac{y}{2} - \frac{\varepsilon}{2}}^{\frac{y}{2} + \frac{\varepsilon}{2}} f_{X_1}(x_1) dx_1 + \int_{y - \varepsilon}^{y + \varepsilon} f_{X_1}(x_1) dx_1} \\ &\approx \lim_{\varepsilon \rightarrow 0} \frac{-\frac{y}{2} f_{X_1}(\frac{y}{2}) \varepsilon + y f_{X_1}(y) 2\varepsilon}{f_{X_1}(\frac{y}{2}) \varepsilon + f_{X_1}(y) 2\varepsilon} \\ &= \frac{-\frac{y}{2} f_{X_1}(\frac{y}{2}) + 2y f_{X_1}(y)}{f_{X_1}(\frac{y}{2}) + 2f_{X_1}(y)}. \end{aligned} \tag{4.7}$$

It is worth noting that this function is in agreement with the function derived by Broome (1995:p10) for the  $E(Z|Y = y)$ . By subtracting the observed amount  $Y = y$  from the expression derived by Broome (1995), the function as derived above can be obtained.

#### 4.3.4 A “halving-only” strategy

For the “halving-only” process

$$C'_y = \bigcup_{\omega_3 \in \{0;1\}} \{\{x_1, 0, \omega_3\} : 2^{\omega_3}(y - \varepsilon) < x_1 \leq 2^{\omega_3}(y + \varepsilon)\}.$$

Noting, from (3.9), that

$$b(\{c\}) = b(\{x_1; 0; \omega_3\}) = \begin{cases} -\frac{1}{2}x_1, & \text{for } \omega_3 = 0; \\ \frac{1}{2}x_1, & \text{for } \omega_3 = 1, \end{cases}$$

and following from (3.1) then

$$\begin{aligned} E(B|Y = y) &\approx \lim_{\varepsilon \rightarrow 0} E(B|y - \varepsilon < Y \leq y + \varepsilon) \\ &= \lim_{\varepsilon \rightarrow 0} E(B|C'_y) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\int_{C'_y} b(\{c\}) f_{X_1}(x_1) dx_1}{\int_{C'_y} f_{X_1}(x_1) dx_1} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{-\frac{1}{2} \int_{y-\varepsilon}^{y+\varepsilon} x_1 f_{X_1}(x_1) dx_1 + \frac{1}{2} \int_{2y-2\varepsilon}^{2y+2\varepsilon} x_1 f_{X_1}(x_1) dx_1}{\int_{y-\varepsilon}^{y+\varepsilon} f_{X_1}(x_1) dx_1 + \int_{2y-2\varepsilon}^{2y+2\varepsilon} f_{X_1}(x_1) dx_1} \\ &\approx \lim_{\varepsilon \rightarrow 0} \frac{-y f_{X_1}(y) \varepsilon + 4y f_{X_1}(y) \varepsilon}{2 f_{X_1}(y) \varepsilon + 4 f_{X_1}(2y) \varepsilon} \\ &= \frac{-y f_{X_1}(y) + 4y f_{X_1}(2y)}{2 f_{X_1}(y) + 4 f_{X_1}(2y)}. \end{aligned} \tag{4.8}$$

#### 4.3.5 A “halving or doubling” strategy

The combined events associated with the “halving-only” and “doubling-only” process are necessary for the “halving or doubling” process.

$$\mathcal{C}_y = C_y \cup C'_y.$$

However, if  $f_{X_1}(x)$  were continuous, the following approach would be adopted. For the agent to observe an amount in the allocated envelope, from the interval  $y - \varepsilon < Y \leq y + \varepsilon$ , he or she would, assuming the “halving or doubling” process, have to consider the activities of the host in drawing an initial amount from the set of intervals:

$$\begin{aligned}
E(B|Y = y) &\approx \lim_{\varepsilon \rightarrow 0} E(B|y - \varepsilon < Y \leq y + \varepsilon) \\
&= \lim_{\varepsilon \rightarrow 0} E(B|\mathcal{C}_y) \\
&= \lim_{\varepsilon \rightarrow 0} E(B|C_y \cup C'_y) \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\int_{C_y} b(\{c\})f_{X_1}(x_1)dx_1 + \int_{C'_y} b(\{c\})f_{X_1}(x_1)dx_1}{\int_{C_y} f_{X_1}(x_1)dx_1 + \int_{C'_y} f_{X_1}(x_1)dx_1} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{-\int_{\frac{y}{2}-\varepsilon}^{\frac{y}{2}+\varepsilon} x_1 f_{X_1}(x_1)dx_1 + \frac{1}{2} \int_{y-\varepsilon}^{y+\varepsilon} x_1 f_{X_1}(x_1)dx_1 + \frac{1}{2} \int_{2y-2\varepsilon}^{2y+2\varepsilon} x_1 f_{X_1}(x_1)dx_1}{\int_{\frac{y}{2}-\varepsilon}^{\frac{y}{2}+\varepsilon} f_{X_1}(x_1)dx_1 + 2 \int_{y-\varepsilon}^{y+\varepsilon} f_{X_1}(x_1)dx_1 + \int_{2y-2\varepsilon}^{2y+2\varepsilon} f_{X_1}(x_1)dx_1} \\
&\approx \lim_{\varepsilon \rightarrow 0} \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2})\varepsilon + \frac{y}{2}f_{X_1}(y)2\varepsilon + yf_{X_1}(2y)4\varepsilon}{f_{X_1}(\frac{y}{2})\varepsilon + 2f_{X_1}(y)2\varepsilon + f_{X_1}(2y)4\varepsilon} \\
&= \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2}) + yf_{X_1}(y) + 4yf_{X_1}(2y)}{f_{X_1}(\frac{y}{2}) + 4f_{X_1}(y) + 4f_{X_1}(2y)}. \tag{4.9}
\end{aligned}$$

Hence, the identification of the content assigned to the envelopes and the subsequent allocation of the primed envelopes is independent of any game strategy identified by the agent. However, the benefit of switching must be conditional upon the realised (observed) amount  $Y = y$ , and not the distinctly different  $Y = X_1 = x_1$  or  $Y = X'_1 = (1/2(1 - \omega_2) + 2\omega_2)X_1$ . The only events that generate  $Y = y$ , the amount observed by the agent in the allocated envelope, are associated with the host selecting an initial amount  $X_1 = x_1$  where  $x_1 \in \{y/2; y; 2y\}$  and assigning this amount to the first envelope.

## 4.4 A switching strategy

If the host were to impose bounds on the amounts allocated, then the agent would need to consider what impact this would have on the host's activities and consequently on his or her own strategy.

The agent, if informed about the density function from which the monetary amounts were

drawn should play as follows. If a lower bound  $x_l$  is specified and all monetary amounts are greater than or equal to this lower bound  $x_l$ , then if the content of the chosen envelope is revealed to be less than  $2x_l$ , the agent should switch. Similarly, if an upper bound  $x_u$  is specified and all monetary amounts are less than or equal to this upper bound  $x_u$ , then if the content is revealed to be greater than  $x_u/2$ , the agent should not switch. For all other observed amounts  $2x_l < y \leq x_u/2$ , the agent would need to consider  $E(B|Y = y)$ . If  $E(B|Y = y) < 0$ , then the agent should not switch, whereas if  $E(B|Y = y) > 0$ , then the agent should switch, otherwise he or she should be indifferent.

The switching strategy  $s(y)$  can be expressed succinctly using an indicator function:

$$s(y) = yI_{[x_l; 2x_l]} + E(B|Y = y)I_{[2x_l; \frac{1}{2}x_u]} - \frac{y}{2}I_{[\frac{1}{2}x_u; x_u]}. \quad (4.10)$$

The strategy for any observed  $x_l \leq y \leq x_u$  is as follows: For  $s(y) < 0$ , the agent should not switch. For  $s(y) > 0$ , the agent should switch, and for  $s(y) = 0$ , the agent should be indifferent.

It will be necessary for the agent to consider what the impact will be at the boundaries  $2x_l$  and  $x_u/2$ . At these transition points it may necessitate that the agent evaluate the expected benefit of switching, before deciding a particular action.

## 4.5 Other strategies

There are two other possible strategies that can be considered. One is associated with the host priming the first envelope and allocating the envelope to the agent before assigning an amount to the second and complementary envelope. And reciprocal to this is when the host primes the first envelope and retains the envelope as the complement. The second envelope is then primed before allocating it to the agent. It will be noted that

the assigning of the amount to the second envelope, in both these instances, must be associated with the “halving or doubling” process.

#### 4.5.1 Allocate the first envelope

A strategy associated with the former is nothing more than a simple lottery in which you have for certain the amount  $Y = y$  and the complementary envelope contains the expected amount  $Z = y/4 + y$ . In this instance there is an obvious benefit  $b(y) = y/4$  to be had for switching. The events of relevance in this instance are  $C_y = \{\{x_1; \omega_2; 0\} : \{y; 0; 0\}; \{y; 1; 0\}\}$ . Here the content and process excludes switching  $\omega_3 = 0$  and the probability of the initial amount is of no relevance in the switching strategy. The benefit

$$\begin{aligned} b(\{c\}) &= b(\{x_1; \omega_2; 0\}) \\ &= \begin{cases} -\frac{1}{2}x_1, & \text{if } \omega_2 = 0; \\ x_1, & \text{if } \omega_2 = 1. \end{cases} \end{aligned}$$

For discrete variables the expected benefit would be

$$\begin{aligned} E(B|Y = y) &= E(B|C_y) \\ &= \frac{\sum_{\{c\} \in C_y} b(\{c\})P(\{c\})}{\sum_{\{c\} \in C_y} P(\{c\})} \\ &= \frac{-\frac{y}{2}P(X_1 = y) + yP(X_1 = y)}{P(X_1 = y) + P(X_1 = y)} \\ &= \frac{y}{4} > 0. \end{aligned} \tag{4.11}$$

While for continuous variables

$$C_y = \{\{x_1; \omega_2; 0\} : y - \varepsilon < x_1 \leq y + \varepsilon\}$$

and following from (3.1) then

$$\begin{aligned}
E(B|Y = y) &\approx \lim_{\varepsilon \rightarrow 0} E(B|y - \varepsilon < Y \leq y + \varepsilon) \\
&= \lim_{\varepsilon \rightarrow 0} E(B|C_y) \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\int_{C_y} b(\{c\}) f_{X_1}(x_1) dx_1}{\int_{C_y} f_{X_1}(x_1) dx_1} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{-\frac{1}{2} \int_{y-\varepsilon}^{y+\varepsilon} x_1 f_{X_1}(x_1) dx_1 + \int_{y-\varepsilon}^{y+\varepsilon} x_1 f_{X_1}(x_1) dx_1}{\int_{y-\varepsilon}^{y+\varepsilon} f_{X_1}(x_1) dx_1 + \int_{y-\varepsilon}^{y+\varepsilon} f_{X_1}(x_1) dx_1} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\frac{1}{2} \int_{y-\varepsilon}^{y+\varepsilon} x_1 f_{X_1}(x_1) dx_1}{2 \int_{y-\varepsilon}^{y+\varepsilon} f_{X_1}(x_1) dx_1} \\
&\approx \lim_{\varepsilon \rightarrow 0} \frac{\frac{1}{2} y f_{X_1}(y) 2\varepsilon}{2 f_{X_1}(y) 2\varepsilon} \\
&= \frac{y}{4} > 0.
\end{aligned} \tag{4.12}$$

and hence, irrespective of the type of variable, the same expected benefit is realised and it would therefore always be beneficial for the agent to switch envelopes.

This strategy is alluded to by Cargile (1992:p212) and amounts considering a gamble on the outcome of a fair coin at payoff of double or half, while in receipt of a certain amount  $Y = y$ .

#### 4.5.2 Allocate the second envelope

A strategy associated with the later is obtained similarly. The agent will reveal the content of the allocated envelope to be  $Y = y$ . In this instance for the host may have



generated an initial amount  $X_1 = y/2$  or  $X_1 = 2y$ . The only possible host activity is associated with the events  $C_y = \{\{x_1; \omega_2; 1\} : \{y/2; 1; 1\}; \{2y; 0; 1\}\}$ . Here the process includes switching  $\omega_3 = 1$  and the probability of the initial amounts  $X_1 = y/2$  and  $X_1 = 2y$  is significant in determining a switching strategy. Any strategy associated with such a content and allocation process is dependent on the distribution function. The benefit would be

$$\begin{aligned} b(\{c\}) &= b(\{x_1; \omega_2; 1\}) \\ &= \begin{cases} \frac{1}{2}x_1, & \text{if } \omega_2 = 0; \\ -x_1, & \text{if } \omega_2 = 1. \end{cases} \end{aligned}$$

For discrete variables the expected benefit would be

$$\begin{aligned} E(B|Y = y) &= E(B|C_y) \\ &= \frac{\sum_{\{c\} \in C_y} b(c)P(c)}{\sum_{\{c\} \in C_y} P(c)} \\ &= \frac{-\frac{y}{2}P(X_1 = \frac{y}{2}) + yP(X_1 = 2y)}{P(X_1 = \frac{y}{2}) + P(X_1 = 2y)}. \end{aligned} \tag{4.13}$$

Switching in this instance is now conditional upon the probability density function and it will be always be beneficial to switch when

$$2P(2y) > P(\frac{y}{2}).$$

While for continuous variables

$$C_y = \{\{x_1; \omega_2; 1\} : 2^{1-2\omega_2}(y - \varepsilon) < x_1 \leq 2^{1-2\omega_2}(y + \varepsilon)\}$$

then

$$\begin{aligned}
E(B|Y = y) &\approx \lim_{\varepsilon \rightarrow 0} E(B|y - \varepsilon < Y \leq y + \varepsilon) \\
&= \lim_{\varepsilon \rightarrow 0} E(B|C_y) \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\int_{C_y} b(\{c\}) f_{X_1}(x_1) dx_1}{\int_{C_y} f_{X_1}(x_1) dx_1} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\frac{1}{2} \int_{2y-2\varepsilon}^{2y+2\varepsilon} x_1 f_{X_1}(x_1) dx_1 - \int_{\frac{y}{2}-\frac{\varepsilon}{2}}^{\frac{y}{2}+\frac{\varepsilon}{2}} x_1 f_{X_1}(x_1) dx_1}{\int_{2y-2\varepsilon}^{2y+2\varepsilon} f_{X_1}(x_1) dx_1 + \int_{\frac{y}{2}-\frac{\varepsilon}{2}}^{\frac{y}{2}+\frac{\varepsilon}{2}} f_{X_1}(x_1) dx_1} \\
&\approx \lim_{\varepsilon \rightarrow 0} \frac{-\frac{y}{2} f_{X_1}(\frac{y}{2}) \varepsilon + 4y f_{X_1}(2y) \varepsilon}{f_{X_1}(\frac{y}{2}) \varepsilon + 4f_{X_1}(2y) \varepsilon} \\
&= \frac{-\frac{y}{2} f_{X_1}(\frac{y}{2}) + 4y f_{X_1}(2y)}{f_{X_1}(\frac{y}{2}) + 4f_{X_1}(2y)}. \tag{4.14}
\end{aligned}$$

Switching in this instance is now conditional upon the probability density function and it will be always be beneficial to switch when

$$8f_{X_1}(2y) > f_{X_1}(\frac{y}{2}).$$

## 4.6 Summary

It is essential for the agent to be informed about the content and allocation process adopted by the host in order to derive an appropriate switching strategy. As we have demonstrated, the envelope-switching strategy is dependent on the content and allocation process adopted by the host, and the distribution function from which the host sampled the initial amount assigned to the first envelope. Once the agent is in possession of this

information, the strategy is easily derived. This information includes not only the sighting of the content of the allocated envelope, but also the adopted content and allocation process.

A switching strategy to optimise the expected outcome for any observed  $Y = y$  is possible if the agent is provided with this information. Moreover, the agent also needs to be informed about the bounds  $x_l$  and  $x_u$  if they are applied. Hence, two distinct strategies are available depending on the content and allocation process. Each strategy, depending on the distribution function from which the initial amount is chosen, may result in a different decision for any observed amount in the allocated envelope.

From the derived strategies it will be realised that the content of the complementary envelope has no impact on the switching strategy. The content of the complementary envelope is of nothing more than curiosity value?

# Chapter 5

## Envelope-switching strategies

Some distribution functions are labelled paradoxical by Broome (1995). Such distributions are by their nature not paradoxical. Such an attribute is only possible because the rate of monetary benefit exceeds the probability of such an event. As demonstrated by the Weibull density function, some distribution functions that are not paradoxical. The choice of density functions to demonstrate the aspects of the two-envelope problem for discrete and continuous random variables is now presented. Many of these have already been explored in the literature and are repeated here.

### 5.1 Uniform distribution

Suppose the agent is informed that the host has chosen to randomly sample the initial amount from the standard uniform probability density function and that

$$f_{X_1}(x_1) = 1 \quad \forall 0 < x_1 \leq 1.$$

Since  $X_1$  is a continuous random variable, applying (4.9) and (3.3) for the halving or doubling process, we find

$$E(B|Y = y) = \frac{-\frac{y}{2} + y + 4y}{1 + 4 + 4} = \frac{y}{2}.$$

Meanwhile, by applying 4.7 and (3.5) to the doubling-only process, we find

$$E(B|Y = y) = \frac{-\frac{y}{2} + 2y}{1 + 2} = \frac{y}{2}.$$

Therefore, for either of the content and allocation processes, the indicator function (4.10) yields the same envelope-switching strategy:

$$s(y) = \begin{cases} \frac{y}{2}, & \text{if } 0 < y \leq \frac{1}{2}, \text{ therefore switch;} \\ -\frac{y}{2}, & \text{if } \frac{1}{2} < y \leq 1, \text{ therefore do not switch.} \end{cases}$$

## 5.2 Weibull distribution

Suppose the agent is informed that the host has chosen to randomly sample the initial amount from a Weibull probability density function and that

$$X \sim f_X(x, \lambda, k) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \quad \forall x \geq 0.$$

We set the parameter

$$\lambda = \frac{1}{2}$$

and for  $k = 2$  the distribution function is also referred to as the “Rayleigh distribution”.

Then

$$X \sim f_X(x) = f_X\left(x, \frac{1}{2}, 2\right) = 8xe^{-(2x)^2} \quad \forall x > 0.$$

The expected benefit for an observed  $Y = y$  is then expressed as follows.

### 5.2.1 Strategy associated with the “doubling-only” process

Applying (4.7)

$$\begin{aligned}
E(B|Y = y) &= \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2}) + 2yf_{X_1}(y)}{f_{X_1}(\frac{y}{2}) + 2f_{X_1}(y)} \\
&= y \frac{8e^{-3y^2} - 1}{8e^{-3y^2} + 2}.
\end{aligned}$$

The concept of an exchange condition  $e(y)$  has been introduced by Blachman and Kilgour (2001:p175) and as a result

$$E(B|Y = y) = \frac{e(y)}{8e^{-3y^2} + 2}.$$

The denominator in  $E(B|Y = y)$  will always be positive. Solving for  $y$  in the exchange condition  $e(y) = 0$ . Then  $8e^{-3y^2} - 1 = 0$  and therefore  $y = \sqrt{\ln 2}$ .

Consequently the switching strategy  $s(y)$ , when observing  $Y = y$ , is

$$s(y) = \begin{cases} \text{Switch,} & \text{if } y < \sqrt{\ln 2}; \\ \text{Indifferent,} & \text{if } y = \sqrt{\ln 2}; \\ \text{Do not switch,} & \text{if } y > \sqrt{\ln 2}. \end{cases}$$

### 5.2.2 Strategy associated with the “halving and doubling” process

Applying (4.9)

$$\begin{aligned}
E(B|Y = y) &= \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2}) + yf_{X_1}(y) + 4yf_{X_1}(2y)}{f_{X_1}(\frac{y}{2}) + 4f_{X_1}(y) + 4f_{X_1}(2y)} \\
&= y \frac{32e^{-15y^2} + 4e^{-3y^2} - 1}{32e^{-15y^2} + 16e^{-3y^2} + 2}
\end{aligned}$$

Solving for  $y$  in the following exchange condition

$$e(y) = 32e^{-15y^2} + 4e^{-3y^2} - 1 = 0.$$

Then

$$y \approx 0.686511 = k.$$

Consequently, for the “halving or doubling” process, the switching strategy  $s(y)$ , when observing  $Y = y$ , is

$$s(y) = \begin{cases} \text{Switch,} & \text{if } y < k; \\ \text{Indifferent,} & \text{if } y = k; \\ \text{Do not switch,} & \text{if } y > k \end{cases}$$

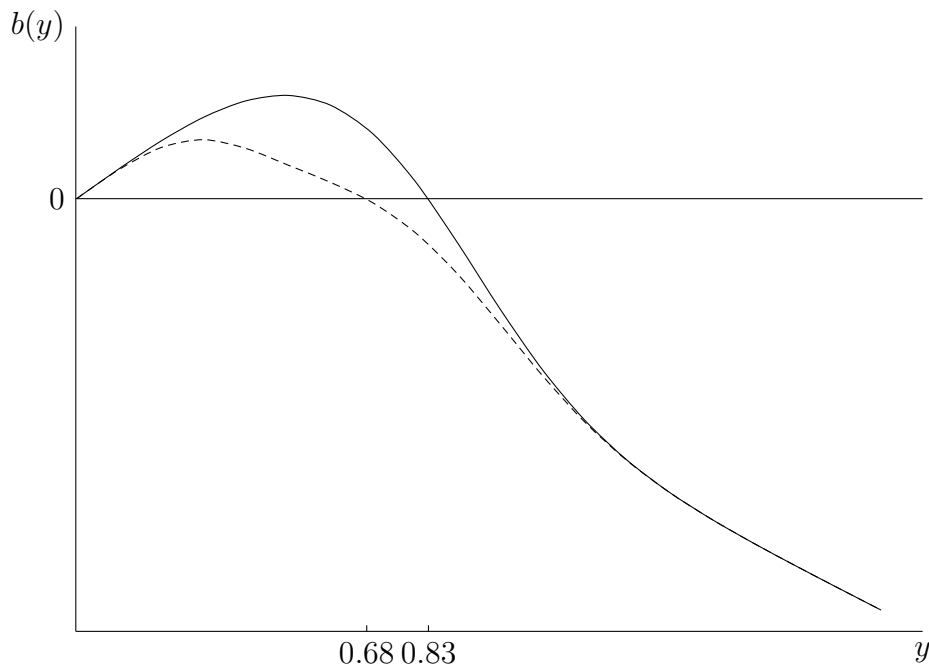


Figure 5.1: Weibull: The expected benefit  $b(0.686511) \approx 0$  when the “halving or doubling” strategy is adopted and  $b(\sqrt{\ln 2}) = 0$  when the “doubling-only” strategy is adopted.

In this example, it is noted that the benefit functions associated with the different content

and allocation processes highlight the difference in strategies. However, both in Figure 5.1 indicate that the expected benefit becomes negative as the observed value increases and that the agent should therefore not switch.

## 5.3 Broome's approach

The following examples, that are described by Broome (1995:p7), are associated with a “doubling-only” content and allocation process.

### 5.3.1 Discrete variable

$$f_{X_1}(x_1) = \frac{2^n}{3^{n+1}} \quad \forall x_1 = 2^n, n \in \mathbb{N}_0.$$

Applying (4.4)

$$\begin{aligned} E(B|Y = y) &= \frac{-\frac{y}{2}P(X_1 = \frac{y}{2}) + yP(X_1 = y)}{P(X_1 = \frac{y}{2}) + P(X_1 = y)} \\ &= y \frac{2^n - 2^{n-2}3}{2^{n-1}3 + 2^n} \\ &= \frac{y}{10} > 0. \end{aligned}$$

And, consequently, it is beneficial to switch irrespective of the amount observed.

### 5.3.2 Continuous variable

$$f_{X_1}(x_1) = \frac{1}{(x_1 + 1)^2} \quad \forall x_1 > 0.$$

Applying (4.7)



$$\begin{aligned}
E(B|Y = y) &= \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2}) + 2yf_{X_1}(y)}{f_{X_1}(\frac{y}{2}) + 2f_{X_1}(y)} \\
&= y \frac{(y+2)^2 - (y+1)^2}{(y+2)^2 + 2(y+1)^2} \\
&= \frac{2y^2 + 3y}{3y^2 + 8y + 6} > 0.
\end{aligned}$$

Therefore, always switch irrespective of the amount observed.

## 5.4 Extreme values

The exchange condition that is proposed by Brams and Kilgour (1995:p28) is developed on the “halving-only” process, since a prior distribution is defined for the larger amount in this instance defined by  $X$  and consequently  $X'_1 = X/2$ .

While Brams and Kilgour (1995) do not distinguish between the initial amounts  $X_1$  and  $X'_1$  and the final amounts  $Y$  and  $Z$  it must be realised that  $Y = y$  is observed as a result of two distinctly and independent events  $c = (x_1; 0; 0)$  or  $c = (2x_1; 0, 1)$ .

### 5.4.1 Discrete variable

$$P(X_1 = x_1) = p_{k+1} \quad \forall x_1 = 2^k m, k \in \mathbb{Z}.$$

Applying (4.5) the expected benefit upon observing  $Y = y$  is

$$\begin{aligned}
E(B|Y = y) &= \frac{-\frac{y}{2}P(X_1 = y) + yP(X_1 = 2y)}{P(X_1 = y) + P(X_1 = 2y)} \\
&= m \cdot 2^k \frac{2p_{k+2} - p_{k+1}}{2p_{k+2} + 2p_{k+1}}.
\end{aligned}$$

The exchange condition is

$$e(y) = 2p_{k+2} - p_{k+1} \quad \forall y = m \cdot 2^k, k \in \mathbb{Z}.$$

And consequently the switching strategy will be

$$s(y) = \begin{cases} \text{Switch,} & \text{if } e(y) > 0 \\ \text{Indifference,} & \text{if } e(y) = 0 \\ \text{Do no switch,} & \text{if } e(y) < 0. \end{cases}$$

#### 5.4.2 Continuous variable

$$f_{X_1}(x_1) = \begin{cases} 0, & \text{if } x_1 < 1; \\ 10^{-2k-1}, & \text{if } 10^k \leq x_1 < 10^{k+1} \quad \forall k \in \mathbb{N}_0. \end{cases}$$

If the content of the allocated envelope is revealed to contain an amount  $1/2 \leq y < 1$  then the agent can take it that such an instance could only have arisen as a result of the host priming the first envelope with an amount  $1 \leq x_1 < 2$  and as a result of the “halving-only” process the second envelope would be assigned an amount  $1/2 \leq x'_1 < 1$ . And it is this second envelope that yields the observed amount. No other events are possible for this outcome. The probability of any event  $c = (x_1, 0, 1)$  for all  $1 \leq x_1 < 2$ , with an associated  $P(c) = 10^{-1}$ , implies that the agent should always switch.

The agent will now need to consider observing an amount from the interval  $10^k \leq y < 10^{k+1}$  for all  $k \in \mathbb{N}_0$ . The agent will need to keep in mind that the amount observed  $Y = y$  is possible only as a result of the activities of the host. It is these activities and the resultant outcomes (events) dictate what is observed by the agent.

The agent will noting that, since the “halving-only” process is operating, consider three possibilities.

**Interval 1.**  $10^k \leq y < 2 \cdot 10^k$

*The agent considers observing an amount from the intervals  $10^k \leq y < 2 \cdot 10^k$  for all  $k \in \mathbb{N}_0$ .*

This can be realised for any event  $\{c\} = \{x_1, 0, 0\}$  for all  $10^k \leq X_1 < 2 \cdot 10^k$  or for any event  $\{c\} = \{x_1, 0, 1\}$  for all  $2 \cdot 10^k \leq X_1 < 4 \cdot 10^k$ . The probability of any of these events is  $P(\{c\}) = 10^{-2k-1}$  for all  $k \in \mathbb{N}_0$ . Applying (4.8)

$$E(B|Y = y) = \frac{-yP(10^k \leq X_1 < 2 \cdot 10^k) + 4yP(2 \cdot 10^k \leq X_1 < 4 \cdot 10^k)}{2P(10^k \leq X_1 < 2 \cdot 10^k) + 4P(2 \cdot 10^k \leq X_1 < 4 \cdot 10^k)}.$$

Since,  $P(10^k \leq X_1 < 2 \cdot 10^k) = P(2 \cdot 10^k \leq X_1 < 4 \cdot 10^k)$  for all  $k \in \mathbb{N}_0$ , then

$$\begin{aligned} &= \frac{-y + 4y}{2 + 4} \\ &= \frac{y}{2} > 0 \quad \forall k \in \mathbb{N}_0, 10^k \leq Y < 2 \cdot 10^k. \end{aligned}$$

**Interval 2.**  $10^{k+1}/2 \leq y < 10^{k+1}$

The agent considers observing an amount from the interval  $10^{k+1}/2 \leq y < 10^{k+1}$ .

This can be realised for any event  $\{c\} = \{x_1, 0, 0\}$   $\forall 10^{k+1}/2 \leq X_1 < 10^{k+1}$  with an associated probability  $P(\{c\}) = 10^{-2k-1}$  for all  $k \in \mathbb{N}_0$  or for any event  $\{c\} = \{x_1, 0, 1\}$   $\forall 10^{k+1} \leq X_1 < 10^{k+1}2$  with an associated probability  $P(\{c\}) = 10^{-2(k+1)-1}$  for all  $k \in \mathbb{N}_0$ . Applying (4.8)

$$\begin{aligned} E(B|Y = y) &= \frac{-yP(\frac{1}{2}10^{k+1} \leq X_1 < 10^{k+1}) + 4yP(10^{k+1} \leq X_1 < 2 \cdot 10^{k+1})}{2P(\frac{1}{2}10^{k+1} \leq X_1 < 10^{k+1}) + 4P(10^{k+1} \leq X_1 < 2 \cdot 10^{k+1})} \\ &= \frac{-y10^{-2k-1} + 4y10^{-2(k+1)-1}}{210^{-2k-1} + 410^{-2(k+1)-1}} \\ &= -y\frac{24}{51} < 0 \quad \forall k \in \mathbb{N}_0, \frac{1}{2}10^{k+1} \leq Y < 10^{k+1}. \end{aligned}$$

**Interval 3.**  $10^k2 \leq y < 10^{k+1}/2$

And finally, the agent considers observing an amount from the interval  $10^k2 \leq y < 10^{k+1}/2$ .

For any event  $\{c\} = \{x_1, 0, 0\}$   $\forall 10^k2 \leq X_1 < 10^{k+1}/2$  with an associated probability  $P(\{c\}) = 10^{-2k-1}$  for all  $k \in \mathbb{N}_0$  or for any event  $\{c\} = (x_1, 0, 1)$   $\forall 10^k4 \leq X_1 < 10^{k+1}$  with an associated probability  $P(\{c\}) = 10^{-2k-1}$  for all  $k \in \mathbb{N}_0$ . Applying (4.8)

$$E(B|Y = y) = \frac{-yP(2 \cdot 10^k2 \leq X_1 < \frac{1}{2}10^{k+1}) + 4yP(4 \cdot 10^k \leq X_1 < 10^{k+1})}{2P(2 \cdot 10^k \leq X_1 < \frac{1}{2}10^{k+1}) + 4P(4 \cdot 10^k \leq X_1 < 10^{k+1})}$$

Since,  $P(2 \cdot 10^k \leq X_1 < \frac{1}{2}10^{k+1}) = P(4 \cdot 10^k \leq X_1 < 10^{k+1})$  for all  $k \in \mathbb{N}_0$ , then

$$\begin{aligned} &= \frac{-y + 4y}{2 + 4} \\ &= \frac{y}{2} > 0 \quad \forall k \in \mathbb{N}_0, 10^k 2 \leq Y < \frac{1}{2}10^{k+1}. \end{aligned}$$

We will now consider the instance where  $k \in \mathbb{N}_0$ . When the agent observes an amount  $10^k \leq Y < 10^{k+1}$  then in the presence of the “halving-only” process, what is observed in this interval is only possible if the amount assigned to the first envelope originated from the interval  $10^k \leq X_1 < 10^{k+1}2$ . What must now be allowed for is the random allocation of the envelopes. One, of which will reveal the content to be in the interval  $10^k \leq Y < 10^{k+1}$ .

If an amount  $10^{k+1} \leq X_1 < 2 \cdot 10^{k+1}$  is assigned to the first envelope, and as a result of halving the amount  $10^{k+1}/2 \leq X'_1 < 10^{k+1}$  is assigned to the second envelope. It is the content of this second envelope that the agent will be sighting.

If an amount  $10^{k+1}/2 \leq X_1 < 10^{k+1}$  is assigned to the first envelope, and as a result of halving the amount  $10^{k+1}/4 \leq X'_1 < 10^{k+1}/2$  is assigned to the second envelope. It is the content of the first envelope that the agent will be sighting.

Alternatively, if the first envelope is assigned an amount  $10^k \leq X_1 < 2 \cdot 10^k$  and the second  $10^k/2 \leq X'_1 < 10^k$ .

And consequently, for all  $k \in \mathbb{N}_0$  the agents strategy will be

$$s(y) = \begin{cases} \text{Switch,} & \text{if } \frac{1}{2} \leq y < 1; \\ \text{Switch,} & \text{if } 10^k \leq y < 2 \cdot 10^k; \\ \text{Switch,} & \text{if } 2 \cdot 10^k \leq Y < \frac{1}{2}10^{k+1}; \\ \text{Do not switch,} & \text{if } \frac{1}{2}10^{k+1} \leq Y < 10^{k+1}. \end{cases}$$

## 5.5 Recurrence relation

Meacham and Weisberg (2003) questions the findings of Clark and Shackel (2000) and propose the following distribution function associated with a discrete variable.

$$p_n = P(X_1 = 2^n) = \frac{1}{2}(p_{n-1} + 2^{-2n}) = \frac{1}{2}p_{n-1} + 2^{-(2n+1)} \quad \forall n \in \mathbb{N}$$

and

$$P(X_1 = 2^0) = p_0 = \frac{1}{12}.$$

Consider

$$p_k = P(X_1 = 2^k) = \frac{1}{2}p_{k-1} + 2^{-(2k+1)}$$

and

$$p_{k+1} = P(X_1 = 2^{k+1}) = \frac{1}{2}p_k + 2^{-(2k+3)}.$$

The “doubling-only” process will be adopted. Applying (4.4)

$$\begin{aligned} E(B|Y = 2^{k+1}) &= \frac{-2^k P(X_1 = 2^k) + 2^{k+1} P(X_1 = 2^{k+1})}{P(X_1 = 2^k) + P(X_1 = 2^{k+1})} \\ &= \frac{-2^k p_k + 2^{k+1} p_{k+1}}{p_k + p_{k+1}} \\ &= \frac{-2^k p_k + 2^{k+1} (\frac{1}{2}p_k + 2^{-(2k+3)})}{p_k + \frac{1}{2}p_k + 2^{-(2k+3)}} \\ &= \frac{2^{-(k+2)}}{\frac{3}{2}p_k + 2^{-(2k+3)}} \\ &= \frac{2^{-(k+1)}}{3p_k + 2^{-2(k+1)}} > 0 \quad \forall k \in \mathbb{N}_0. \end{aligned}$$

And consequently it is beneficial to switch irrespective of the amount observed.

## 5.6 Improper distribution

The mere idea of considering a distribution function that is not normalised would be anathema to many of the authors. As had been documented, only normalised distributions are considered by many authors and used to explain the apparent paradox. At the outset, to consider such distributions may seem unacceptable (Vazquez 2012:p11). However, the expected benefit conditional on the sub- $\sigma$ -algebra generated by the observed amount  $Y = y$  for a improper distributions are now considered.

### 5.6.1 Exponential

For illustrative purposes consider

$$f_{X_1}(x_1) = 2^{-4x^2} \quad \forall x > 0.$$

Although this function possess similarities to the Weibull it is not normalised and therefore a constant  $k$  will be introduced to ensure that it satisfies the criteria of a probability density function (cumulative distribution function).

The probability density function is therefore

$$f_{X_1}(x_1) = k2^{-4x^2} \quad \forall x > 0.$$

A “doubling-only” process will be adopted, and since  $X_1$  is a continuous variable, then applying (4.7)

$$\begin{aligned}
E(B|Y = y) &= \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2}) + 2yf_{X_1}(y)}{f_{X_1}(\frac{y}{2}) + 2f_{X_1}(y)} \\
&= \frac{-\frac{y}{2}k2^{-y^2} + 2yk2^{-4y^2}}{k2^{-y^2} + 2k2^{-4y^2}}.
\end{aligned}$$

It will be noted that the normalising constant  $k$  is factored out and that the exchange condition is

$$e(y) = -\frac{y}{2}2^{-y^2} + 2y2^{-4y^2}.$$

From which the strategy associated with the function is as follows:

$$s(y) = \begin{cases} \text{Switch,} & \text{if } y < \sqrt{\frac{2}{3}}; \\ \text{Indifferent,} & \text{if } y = \sqrt{\frac{2}{3}}; \\ \text{Do not switch,} & \text{if } y > \sqrt{\frac{2}{3}}. \end{cases}$$

And as expected, this function presents a similar switching strategy to that derived for the Weibull distribution.

### 5.6.2 Jeffreys' prior

The scale invariant “Jeffreys' prior”

$$f_{X_1}(x_1) \propto \frac{1}{x_1}$$

is proposed by Vazquez (2012:p7) and considered by Gill (2011:p9).

For illustrative purposes consider

$$f_{X_1}(x_1) = \frac{k}{x_1^n} \quad \forall x > 0, n \in \mathbb{N}.$$

Although there may be protests from the purists, as before a constant  $k$  will be introduced to ensure that it satisfies the criteria of a probability density function (cumulative distribution function).

A “doubling-only” process will be adopted, and since  $X_1$  is a continuous variable, then applying (4.7)

$$\begin{aligned} E(B|Y = y) &= \frac{-\frac{y}{2}f_{X_1}(\frac{y}{2}) + 2yf_{X_1}(y)}{f_{X_1}(\frac{y}{2}) + 2f_{X_1}(y)} \\ &= y \frac{2 - 2^{n-1}}{2 + 2^n}. \end{aligned}$$

It will be noted that the normalising constant  $k$  is factored out, as is  $y^{-n}$ , and that the exchange condition is

$$e(y) = 2 - 2^{n-1}.$$

From which the strategy associated with the function, irrespective of the observed amount, is as follows:

$$s(y) = \begin{cases} \text{Switch,} & \text{if } n = 1; \\ \text{Indifferent,} & \text{if } n = 2; \\ \text{Do not switch,} & \text{if } n > 2. \end{cases}$$

The very idea of an improper distribution may be preposterous to many purists (Scott and Scott 1997:P39). While such distribution functions as noted by Vazquez (2012:p11) may seem absurd, they are, nonetheless, as entertaining as the two-envelope problem and should therefore be considered as legitimate.



## 5.7 Summary

These examples are all legitimate distribution functions and their use to generate an initial amount, although in question by some, do give rise to some interesting and diverse switching strategies. All these strategies are conditional upon the sighting of the amount in the allocated envelope. And essential to any one of the strategies is the content and allocation process adopted by the host.

# Chapter 6

## The question of a “paradox”

... thought is the weighing of relative likelihoods of possible events and the act of sampling from the “posterior”, the probability distribution on unknown events, given the sum total of our knowledge of past events and the present context. If this is so, then the paradigmatic mental object is not a proposition, standing in all its eternal glory with its truth value emblazoned on its chest, but the random variable  $x$ , its value subject to probabilities but still not fixed .... The simplest example where human thinking is clearly of this kind may well be the case where probabilities can be made explicit: gambling. Here we are quite conscious that we are weighing likelihoods (and even calculating them if we are mathematically inclined). If we accept this, the division of mathematics corresponding to this realm of experience is not logic but probability and statistics.

— The Dawning of the Age of Stochasticity (Mumford 1999:5)

The two-envelope paradox has generated much debate within academic circles from diverse disciplines such as mathematics, statistics, economics, philosophy, and psychology. The perception of this research is that irrespective of which envelope is allocated to the agent, it would always be beneficial for the agent to switch to the complementary envelope.

Although Smullyan (1992:191) dismisses the use of probability to explain the paradox, he still uses the words: “probability”, “equal probability”, “chances”, “... the chances of gaining or losing are equal”, and “... odds”. All these terms and expressions have probabilistic connotations and, subsequently, what is considered to be devoid of probability must be resolved by embracing a probabilistic argument. Sutton (2010), by contrast, argues that since the two-envelope problem is a decision theoretic, the agent may assume — in the absence of sighting the content of the allocated envelope — that since the contents are fixed at the time of selection, that they remain unchanged during the process and, therefore, that it is irrelevant whether he or she chooses to switch or not. Whether sighting the content provides any information to assist the agent in deciding to switch or not is dependent on the information provided by the host of the game.

It is true that the content of one envelope is twice that of the other. It is also true that relative to the content of the first envelope — if the halving or doubling process was adopted by the host — the second envelope is either half or double that of the first. However, based on what has been discussed in the previous chapters, this can be considered to be only half the story. The allocation of either envelope to the agent is reliant on the outcome of a random event, and in this instance, there are two possibilities: either the event  $(x_1; \omega_2; 0)$  or the event  $(x_1; \omega_2; 1)$  has been realised by the host. These events may not have occurred with equal likelihood as revealed in Chapter 4.

## 6.1 A likely outcome

“Probability is now irrelevant,” said the Sorcerer.

— Satan, Cantor, and Infinity (Smullyan 1992:191)

Not so! We say that the outcome of a toss of an unbiased coin is either a head or a tail. It is also said that the content of the other envelope is either  $2n$  or  $n/2$ . As stated by

Mumford (1999) such events are best described in the terms of probability.

Before explaining the reasoning behind the apparent paradox, it is necessary to introduce some notation. Where possible, all arguments are expressed in terms of the realisations of the experiments performed by the host; consequently, random variables are only used when necessary in order to align with the thinking of Smullyan and to argue in favour of probability.

The amount assigned to the second envelope is dependent upon the outcome of an experiment performed by the host. The factor  $g(\omega_2)$  is derived from the outcome of tossing an unbiased coin:

$$g(\omega_2) = \frac{1}{2}(1 - \omega_2) + 2\omega_2 = \begin{cases} \frac{1}{2}, & \text{if } \omega_2 = 0; \\ 2, & \text{if } \omega_2 = 1. \end{cases}$$

The random selection of one of the primed (indistinguishable) envelopes and the allocation of that envelope to the agent is represented by the symmetric matrix  $\mathbf{A}(\omega_3)$ .

$$\mathbf{A}(\omega_3) = \begin{bmatrix} 1 - \omega_3 & \omega_3 \\ \omega_3 & 1 - \omega_3 \end{bmatrix} = \begin{cases} I, & \text{if } \omega_3 = 0; \\ J, & \text{if } \omega_3 = 1. \end{cases}$$

where  $I$  and  $J$  are the identity and exchange matrices, respectively.

For the outcomes of the experiments performed by the host, namely  $c = (x_1; \omega_2; 0)$  — allocate the first envelope — and  $c = (x_1; \omega_2; 1)$  — allocate the second envelope, the allocated and complementary envelope contents are as follows:

$$\mathbf{C}(c) = \mathbf{C}(x_1; \omega_2; \omega_3) \tag{6.1}$$

$$\begin{aligned} &= x_1 \begin{bmatrix} 1 & g(\omega_2) \end{bmatrix} \mathbf{A}(\omega_3) \\ &= \begin{cases} x_1 \begin{bmatrix} 1 & g(\omega_2) \end{bmatrix}, & \text{if } \omega_3 = 0; \\ x_1 \begin{bmatrix} g(\omega_2) & 1 \end{bmatrix}, & \text{if } \omega_3 = 1. \end{cases} \end{aligned} \tag{6.2}$$

Therefore, the benefit of switching, given the event  $(x_1; \omega_2; \omega_3)$ , is

$$b(x_1; \omega_2; \omega_3) = \begin{cases} x_1(g(\omega_2) - 1), & \text{if } \omega_3 = 0; \\ x_1(1 - g(\omega_2)), & \text{if } \omega_3 = 1. \end{cases} \tag{6.3}$$

## 6.2 The Story

There are two sealed envelopes on the table. You are told that one of them contains twice as much money as the other. You pick one of them and open it to see how much money is inside. Let's say you find \$100 in it. Then you are given the option of keeping it or trading it for the other envelope. Now, the other envelope has twice as much or half as much with equal probability. Thus, the chances are equal that the other envelope has \$200 or \$50, and the chances of your gaining or losing are equal. And so the odds are in your favour if you trade.

— Satan, Cantor, and Infinity (Smullyan 1992:189)

As mentioned in the literary review the importance of posing the two-envelope problem in a form that supports a probabilistic argument is pivotal for the derivation of a switching strategy. The propositions presented by Smullyan (1992:190) are repeated here for

convenience:

**Proposition 1.** *The amount that you will gain, if you do gain, is greater than the amount you will lose, if you do lose.*

**Proposition 2.** *The amounts are the same.*

For the event  $(x_1; \omega_2; 0)$ , the agent would be allocated the envelope containing the amount  $x_1$ , while the complementary envelope would contain the amount  $x_1 g(\omega_2)$ . Consequently, if the agent were to switch envelopes, the benefit would be  $x_1(g(\omega_2) - 1)$ . The agent should argue similarly for the event  $(x_1; \omega_2; 1)$ .

It is clear from Equation 6.3, in the absence of any information, that  $b(x_1; \omega_2; 0) = -b(x_1; \omega_2; 1)$  supports Proposition 2. In the words of Smullyan (1992:191): “If you gain on the trade, you gain  $d$  dollars, and if you lose on the trade, you lose  $d$  dollars. And so the amounts are the same after all”.

This is akin to the agent having no information about the game and consequently includes the instance where the agent is not afforded the opportunity of sighting the content of the allocated envelope before deciding to switch envelopes or not.

## 6.3 Half the story

Smullyan (1992:p190) has in Proposition 2 chosen to present the equally likely events  $(x_1; \omega_2; 0)$  and  $(x_1; \omega_2; 1)$  and then demonstrates the outcome by way of an example that the benefits of trading amounts are the same.

Suppose, however that the host identifies an amount  $x_1$  and after assigning this amount to the first envelope immediately allocates the envelope to the agent. The host then identifies a subsequent amount  $x_1 g(\omega_2)$  and then assigns this to the second (complementary) envelope.

The only possible event representing the outcome in this instance is  $(x_1; \omega_2; 0)$ , which reflects the allocation to the agent of the envelope that contains  $x_1$ . The complementary envelope will then contain  $x_1 g(\omega_2)$ , which can take the value  $x_1/2$  or  $2x_1$ , depending on  $\omega_2$ . The “halving or doubling” process is in operation here, since a “doubling-only” process would result in  $2x_1$  only and consequently there is no gamble. In other words, the agent knows for certain that the complementary envelope contains twice the amount of the allocated envelope, and therefore a “doubling-only” process is not possible.

For the event  $c = (x_1; \omega_2; 0)$ ,  $x_1$  is allocated to the agent, while the amount in the complementary envelope could be either  $x_1/2$  or  $2x_1$  with an equal likelihood. Hence, if the agent were to switch, he or she would gain  $x_1$  or lose  $x_1/2$ . This is the reasoning behind Proposition 1. It can be argued that a rational agent would switch envelopes if the certain amount  $x_1$  was considered against the pay-off associated with a bet on either  $x_1/2$  or  $2x_1$  (Cargile 1992: p. 212, 216).

A risk-neutral and rational agent should argue that for any gamble that the average payoff of the complementary envelope being greater than the allocated envelope supports the decision to trade. However, in this way no consideration has been paid to the content and allocation process, or more specifically the random selection and allocation of the envelopes. However, to trade and then resort to the same argument is fallacious, since the resultant envelopes amounts to the realisation of the event  $(x_1; \omega_2; 1)$ , and the observed content of the second envelope is now not the amount  $x_1 g(\omega_2)$ , which is a nebulous amount, but either  $x_1/2$  or  $2x_1$ . Neither of these amounts can be observed simultaneously. Thus, had the agent observed  $x_1/2$ , he or she would infer that the complementary envelope contained either  $x_1/4$  or  $x_1$ . In the absence of information, whether having sight of the contents or not, the agent should be indifferent to switching envelopes or not. The benefit of switching, given the event  $(x_1; \omega_2; 0)$ , is therefore

$$b(x_1; \omega_2; 0) = x_1(g(\omega_2) - 1) = \begin{cases} -\frac{1}{2}x_1, & \text{if } \omega_2 = 0; \\ x_1, & \text{if } \omega_2 = 1. \end{cases} \quad (6.4)$$

and since these outcomes are equally likely, the agent should switch envelopes.

## 6.4 The full story

Smullyan (1992) allows for the events  $(x_1; \omega_2; 0)$  and  $(x_1; \omega_2; 1)$  when discussing Proposition 2, but only the event  $(x_1; \omega_2; 0)$  when discussing Proposition 1. Ignoring the event  $(x_1; \omega_2; 1)$  when discussing Proposition 1 amounts to dismissing the random allocation of the envelopes and, consequently, the problem becomes a gamble for which the payoff is the difference between the expected content of the complementary envelope and the sighted content of the allocated envelope.

Possibility 4 and Possibility 5 Before the agent is afforded the opportunity of sighting the content of the allocated envelopes, he or she should consider that the events  $(x_1; \omega_2; 0)$  and  $(x_1; \omega_2; 1)$  are equally likely. For the event  $(x_1; \omega_2; 0)$ , the agent is allocated the envelope that contains  $x_1$  and the complementary envelope therefore contains either  $x_1/2$  or  $2x_1$  with an equal probability. The agent can, in this instance, after sighting the content (Clark and Shackel 2000:428), correctly argue that the complementary envelope contains either half or double the content of allocated envelope. Unfortunately, the same reasoning cannot be applied to the event  $(x_1; \omega_2; 1)$ . In this instance, the content of the allocated envelope will be either  $x_1/2$  or  $2x_1$  and the complementary envelope will contain  $x_1$ . Upon sighting the content of the allocated envelope, if it were  $x_1/2$  the agent would erroneously reason that the content of the complementary envelope would be either  $x_1/4$  or  $x_1$ , of which only  $x_1$  is possible. Similarly, if the agent were to sight the amount  $2x_1$ ,



then using the same reasoning he or she would expect the complementary envelope to contain either  $x_1$  or  $4x_1$ , of which only  $x_1$  is correct.

The switching decision is not dependent on the expected content of the allocated envelope but rather on the sighted content of the allocated envelope and the expected content of the complementary envelope. It is therefore necessary for the agent — upon sighting the content — to apply one of the strategies discussed in Chapter 4.

Moreover, the consideration by Chase (2002:158) of equally likely events  $(x_1; \omega_2; 0)$  and  $(x_1; \omega_2; 1)$  is necessary when comparing Proposition 1 with Proposition 2.

## 6.5 Summary

There is an ample supply of misleading leads, false clues, and dead ends to confuse the issue ... you will be dealing with concepts that are central to a number of philosophical doctrines or controversies. Not only are the fundamental notions of decision theory involved, — preferences, desirability, probability, possibility, and expected utility — but one also finds the slightly suspect Kripkean idea of fixing the reference of a name by means of a description.

— The Mystery of Julius: A Paradox in Decision Theory (Chihara 1995:5)

The introduction of the concept of a probability triple  $(\Omega, \mathcal{F}, P)$  and the defined sample space follows the approach put forward by Christensen and Utts (1992), who explore the paradox from a frequentist perspective. In their proposal, they adopt only the doubling strategy. Bruss (1996) acknowledges that the paradox is a fallacy associated with the failure to define the appropriate sample space. He concurs with Christensen and Utts (1992) in their frequentist and Bayesian explanation.

It is not correct to argue that the complementary envelope  $Z$  contains  $y/2$  or  $2y$  in the absence of the stated information. The events that originate from the activities of the

host are essential for the agent to know. In the absence of this information, sighting the content or not has no bearing on the expected outcome and, consequently, the agent should be indifferent about switching or not.

A fundamental principle for the resolution of any problem is its re-formulation into a format that lends itself readily to the application of methods that may realise an outcome that was not initially obvious. In the absence of such an initiative, academic sleight-of-hand has gone unnoticed. There has thus been the realisation that there is more to this simple problem than meets the eye.

The random variable  $Y$ , which represents the content of the allocated envelope according to current and previous thinking, will always have the value  $Y = y$  and should be treated as fixed.

Smullyan (1992) attempts to explain the paradox using logic; however in Proposition 1 he fails to consider that he was not considering the benefit as he described in Proposition 2 and, therefore, that any comparison would be impossible. However, had he considered expressing Proposition 1 as expressed herein, he would have realised that Proposition 1 and Proposition 2 actually agree.

There are two possibilities. The agent is allocated the first envelope containing the initial amount  $X_1 = x_1$  and the second envelope is then assigned the amount  $X'_1 = x_1/2$  or the agent is allocated the first envelope containing the initial amount  $X_1 = x_1$  and the second envelope is then assigned the amount  $X'_1 = 2x_1$ .

Whether sighting the content or not, in the absence of all other information, Proposition 2 is true. In sighting the content of the allocated envelope, Proposition 1 is not true since only the certain outcome of a specific event is considered.

# Chapter 7

## Conclusion

Today's scientists have substituted mathematics for experiments, and they wander off through equation after equation, and eventually build a structure which has no relation to reality.

— Nikola Tesla

Does the two-envelope paradox exist? This question, which has entertained, challenged, and frustrated many researchers, has received much attention since the original formulation by Kraitchik (1942:p133). In all its variants, however, it remains essentially the same, and therefore the overriding question remains: Is there a benefit to switching? While a number of approaches have proposed to resolve this dilemma, the approach adopted in this paper has revealed a missing piece of the puzzle.

The apparent two-envelope paradox exists only when the problem is formulated from an incomplete or inadequate description of the realisable benefits and when there is complete disregard for some of the fundamental and essential issues associated with probability. For example, Norton (1998) formulates elaborate solutions to a simple problem, and then ignores the fundamental concepts essential for any mathematically derived solution. In conclusion, provided there is no knowledge of the distribution function from which the monetary amounts were drawn or the limits of the monetary content in each of the

envelopes, and irrespective of whether the content of the chosen envelope is sighted or not, there is no benefit to the agent of switching.

This thesis has demonstrated that the benefit of switching therefore does not rely upon a single event but rather upon the statistical expectation associated with the activities of the host for the sighted content of the allocated envelope. Because the posing of the problem is set in probabilistic formalism, it lends itself readily to statistical and probabilistic resolution methods. The activities of the host, when revealed to the agent, are sufficient for the agent to derive an envelope-switching strategy. In the absence of this information, the agent should be indifferent to switching.

Much of what is purported to resolve the problem (and thereby the paradox) results from poor formulation and incorrect structuring as well as from the outright ignorance of fundamental statistical methods. Oftentimes, the problem is contextualised in unnecessary jargon and complex mathematical methods. The paradox that has long possessed academics has no legitimacy, and only exists for those that choose it to. Indeed, the substantiation of the paradox stems from the flawed interpretation of an ill-posed problem and an abuse of basic statistical methods.

The examples, where sighting the content of the allocated envelope, have revealed that in some instances it is always beneficial to switch. There are instances where it is not beneficial to switch and there are others where the strategy is mixed. In all instances the expected benefit is conditional on the sub- $\sigma$ -algebra induced by the observed amount represented by the random variable  $Y$ , irrespective if the observed amount is a probability zero event. Arguments questioning the relevance of distributions or functions over infinite amounts are no longer valid. Distribution or functions that represent infinite amounts are legitimate and in the presence of the necessary information it will be realised that the approach derived in this paper, that exploits the features of the sub- $\sigma$ -algebra, can yield an appropriate switching strategy. As noted by Meacham and Weisberg (2003:p4)

switching envelopes is conditional upon the sighting of a particular value, in the allocated envelope, and not whether swapping is better “on average” over repeated trials. In instances where the “exchange condition” is positive for all observed amounts, implying that it is not necessary to view the contents and that it would make sense to switch envelopes irrespective of sighting the content or not does not legitimise the paradox (Brams and Kilgour 1995:p28).

## 7.1 In closing

... statistical information on frequencies within a large, repetitive class of events is strictly irrelevant to a decision whose outcome depends on a single trial.

— Risk, Ambiguity, and the Savage Axioms (Ellsberg 1961:2)

In the “halving or doubling” process the traditional argument is to assess the benefit of switching the allocated envelope for the complementary envelope. The content of one of the envelopes, for this process, is acknowledged to be either half or double the other. With this information the agent is assumed to reason that there are two amounts for consideration and that the conditional probability of these amounts are all that are necessary to derived the expected benefit. But, as has been demonstrated the conditional expectation for this content and allocation process is dependent on the host sampling three distinct amounts. These possible initial amounts are all induced by the amount observed by the agent.

For any play of the game the agent may lose or the agent may win. Alternatively, if the agent is mathematically inclined, he or she may apply likelihoods to each of the induced outcomes and derive an expected benefit of switching envelopes.

## 7.2 The paradox ...

This is your last chance. After this, there is no turning back. You take the blue pill – the story ends, you wake up in your bed and believe whatever you want to believe. You take the red pill – you stay in Wonderland, and I show you how deep the rabbit hole goes.

— Morpheus (The Matrix)

# Reference

- C. J. Albers, B. P. Kooi, and W. Schaafsma. Trying to resolve the Two-Envelope problem. *Synthese*, 145(1):89–109, May 2005. doi: 10.1007/s11229-004-7665-5. URL <http://dx.doi.org/10.1007/s11229-004-7665-5>. 20, 21
- R. B. Ash. *Real Analysis and Probability*. Academic Pr, June 1972. ISBN 0120652404. 17, 22, 38, 39, 46, 47
- K. B. Athreya and S. N. Lahiri. *Measure Theory and Probability Theory*. Springer, softcover reprint of hardcover 1st ed. 2006 edition, November 2010. ISBN 1441921915. 29
- Nelson M. Blachman and D. Marc Kilgour. Elusive optimality in the box problem. *Mathematics Magazine*, 74(3):pp. 171–181, 2001. ISSN 0025570X. URL <http://0-www.jstor.org.oasis.unisa.ac.za/stable/2690718>. 59
- S. J. Brams and D. M. Kilgour. The box problem: To switch or not to switch. *Mathematics Magazine*, 68(1):27–34, 1995. ISSN 0025570X. URL <http://www.jstor.org/stable/2691373>. 17, 18, 46, 62, 82
- J. Broome. The two-envelope paradox. *Analysis*, 55(1):6–11, 1995. ISSN 00032638. URL <http://www.jstor.org/stable/3328613>. 17, 18, 22, 23, 24, 25, 27, 46, 48, 57, 61
- F. T. Bruss. The fallacy of the two envelopes problem. *Math. Sci.*, 21(2):112–119, 1996. ISSN 0312-3685. 18, 27, 78



- F. T. Bruss and L. Ruschendorf. The switching problem and conditionally specified distribution. *Math. Sci.*, 25(1):47–53, 2000. ISSN 0312-3685. 17, 24, 27
- J. Cargile. On a problem about probability and decision. *Analysis*, 52(4):211–216, 1992. URL <http://www.jstor.org/stable/3328337.html>. 18, 21, 53, 76
- P. Castell and D. Batens. The two envelope paradox: The infinite case. *Analysis*, 54(1):pp. 46–49, 1994. ISSN 00032638. URL <http://0-www.jstor.org.oasis.unisa.ac.za/stable/3328103>. 17
- J. Chase. The non-probabilistic two envelope paradox. *Analysis*, 62(2):157–160, 2002. doi: 10.1093/analys/62.2.157. URL <http://analysis.oxfordjournals.org>. 20, 24, 78
- G. J. Chen. The puzzle of the Two-Envelope puzzle. *SSRN eLibrary*, February 2007. URL [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1132506](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1132506). 21, 24
- C. S. Chihara. The mystery of Julius: A paradox in decision theory. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 80(1):1–16, October 1995. URL <http://www.jstor.org/stable/4320617.html>. 15, 20, 21, 78
- R. Christensen and J. Utts. Bayesian resolution of the “Exchange Paradox”. *The American Statistician*, 46(4):274–276, 1992. ISSN 00031305. URL <http://www.jstor.org/stable/2685310>. 17, 78
- M. Clark and N. Shackel. The two-envelope paradox. *Mind*, 109(435):415–442, 2000. doi: 10.1093/mind/109.435.415. URL <http://mind.oxfordjournals.org/cgi/content/abstract/109/435/415>. 17, 66, 77
- A. de Saint-Exupéry. *The Little Prince*. Egmont UK limited, 1945. 26, 36

- D. Ellsberg. Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669, November 1961. ISSN 00335533. doi: 10.2307/1884324. URL <http://www.jstor.org/discover/10.2307/1884324?uid=2129&uid=2&uid=70&uid=4&sid=21101> ArticleType: primary\_article / Full publication date: Nov., 1961 / Copyright 1961 The MIT Press. 1, 82
- R. D. Gill. *Anna Karenina and The Two Envelopes Problem (Fourth, still very incomplete, draft)*. 2011. URL <http://www.math.leidenuniv.nl/~gill/tep.pdf>. 24, 68
- F. Jackson, P. Menzies, and G. Oppy. The two envelope “Paradox”. *Analysis*, 54(1): 43–45, 1994. 22
- R. C. Jeffrey. Judgmental probability and objective chance. *Erkenntnis*, 24(1):5–16, 1986. ISSN 0165-0106. doi: 10.1007/BF00183197. URL <http://dx.doi.org/10.1007/BF00183197>. 17
- B. D. Katz and D. Olin. A tale of two envelopes. *Mind*, 116(464):pp. 903–926, 2007. ISSN 00264423. URL <http://0-www.jstor.org.oasis.unisa.ac.za/stable/30166515>. 19
- M. Kraitchik. *Mathematical Recreations*. W. W. Norton & Company, Inc., New York (2nd revised edition, 2006, Dover Publications, New York), 1942. ISBN 0-486-45358-8. 1, 80
- T. Lored. Critical thinking, summer 2004: The two-envelope paradox. 2004. URL <http://www.astro.cornell.edu/staff/loredo/bayes/two-envelope.pdf>. 10, 15, 17, 18
- M. D. McDonnell and D. Abbott. Randomized switching in the two-envelope problem. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 2009. doi: 10.1098/rspa.2009.0312. URL

<http://rspa.royalsocietypublishing.org/content/early/2009/07/31/rspa.2009.0312.abst>

24

T. J. McGrew, D. Shier, and H. S. Silverstein. The two-envelope paradox resolved.

*Analysis*, 57(1):28–33, 1997. 17, 24

C. Meacham and J. Weisberg. Clark and Shackel on the Two-Envelope paradox.

[http://mind.oxfordjournals.org/cgi/pdf\\_extract/112/448/685](http://mind.oxfordjournals.org/cgi/pdf_extract/112/448/685), October 2003. URL

[http://mind.oxfordjournals.org/cgi/pdf\\_extract/112/448/685](http://mind.oxfordjournals.org/cgi/pdf_extract/112/448/685). 66, 81

A. M. Mood, F. A. Graybill, and D. C. Boes. *Introduction to the theory of statistics*.

McGraw-Hill, third edition, 1974. ISBN 0-07-042864-6. 46

D. Mumford. The dawning of the age of stochasticity, 1999. URL

<http://www.dam.brown.edu/people/mumford/Papers/OverviewPapers/DawningAgeStoch.pdf>.

71, 73

B. Nalebuff. The other person’s envelope is always greener. *Jour-*

*nal of Economic Perspectives*, 3(1):171–81, Winter 1989. URL

<http://ideas.repec.org/a/aea/jecper/v3y1989i1p171-81.html>. 17, 18

J. Norton. Where the sum of our expectation fails us: The exchange paradox. *Pacific*

*Philosophical Quarterly*, 79(1):34–58, 1998. ISSN 1468-0114. doi: 10.1111/1468-0114.

00049. URL <http://dx.doi.org/10.1111/1468-0114.00049>. 80

E.A. Ok. *Real analysis with economic applications*. Princeton University Press, 2007.

URL <https://files.nyu.edu/eo1/public/Book-PDF/pChapterIII.pdf>. 22, 39, 45,

47

G. Priest and G. Restall. Envelopes and indifference, 2007. URL

<http://philpapers.org/rec/PRIEAI>. 24

- P. Rawling. A note on the two envelopes problem. *Theory and Decision*, 36(1):97–102, 1994. ISSN 0040-5833. doi: 10.1007/BF01075299. URL <http://dx.doi.org/10.1007/BF01075299>. 15, 16, 17, 18, 21, 22, 24
- A. Rényi. *Foundations of Probability*. San Francisco: Holden-Day, 1970. 47
- A. D. Scott and M. Scott. What’s in the two envelope paradox? *Analysis*, 57(1):pp. 34–41, 1997. ISSN 00032638. URL <http://0-www.jstor.org.oasis.unisa.ac.za/stable/3328432>. 17, 69
- E. A. Silver. On mathematical problem posing. *For the Learning of Mathematics*, 14(1):pp. 19–28, 1994. ISSN 02280671. URL <http://0-www.jstor.org.oasis.unisa.ac.za/stable/40248099>. 10, 16
- R. M. Smullyan. *Satan, Cantor, And Infinity and Other Mind-Boggling Puzzles*. Knopf, first printing edition, November 1992. ISBN 0679406883. 20, 21, 72, 74, 75, 77, 79
- J. H. Sobel. Two envelopes. *Theory and Decision*, 36(1):69–96, 1994. 7, 14, 16, 18, 27
- A. Spanos. The Two Envelope Problem: a Paradox or Fallacious Reasoning? *ArXiv e-prints*, January 2013. 22, 24, 38
- P. A. Sutton. The epoch of incredulity: A response to Katz and Olin’s ‘A Tale of Two Envelopes’. *Mind*, 119(473):159–169, 2010. doi: 10.1093/mind/fzp164. URL <http://mind.oxfordjournals.org/content/119/473/159.abstract>. 21, 24, 25, 72
- R. A. Vazquez. The two envelopes probability paradox: Much ado about nothing. *ArXiv e-prints*, June 2012. 21, 67, 68, 69